物联网工程技术及其应用系列规划教材



无线通信原理

WUXIAN TONGXIN YUANLI

主 编 许晓丽 赵明涛

- 公式推导和概念详尽，理论性好
- 阅读材料与案例丰富，可读性强





说　明

本书电子版仅提供给高校任课教师使用，如有任课教师需要全本教材浏览或需要本书课件等相关教学资料，请联系北京大学出版社客服，微信手机同号：15600139606，扫下面二维码可直接联系。

由于教材版权所限，仅限任课教师索取，谢谢！



物联网工程技术及其应用系列规划教材

无线通信原理

主　编　许晓丽　赵明涛

副主编　董光辉　刘靖宇



内 容 简 介

全书共有 8 章，系统地阐述了无线通信领域涉及的基本理论、技术和原理，主要内容包括绪论、无线信道传输、无线通信编码技术、无线调制技术、多址接入技术、无线通信协议、蜂窝移动通信、无线通信新技术。本书以讲透概念、重视方法原理、突出应用和好教易学为指导思想，每章开篇均由相关案例导入本章的学习内容，在理论联系实际的同时，更突显了各章的知识特色。本书还嵌入了很多例题和实际应用的例子，增强了知识的可读性，降低了学习难度。

本书可以作为通信、电子信息专业本科生教材，也可作为通信技术人员的参考书。

图书在版编目(CIP)数据

无线通信原理/许晓丽，赵明涛主编. —北京：北京大学出版社，2014.3

(物联网工程技术及其应用系列规划教材)

ISBN 978-7-301-23705-2

Ⅰ. ①无…　Ⅱ. ①许…②赵…　Ⅲ. ①无线电通信—高等学校—教材　Ⅳ. ①TN92

中国版本图书馆 CIP 数据核字(2014)第 004117 号

书　　　名：无线通信原理

著作责任者：许晓丽　赵明涛　主编

策 划 编 辑：程志强

责 任 编 辑：程志强

标 准 书 号：ISBN 978-7-301-23705-2/TN · 0107

出 版 发 行：北京大学出版社

地　　　址：北京市海淀区成府路 205 号　100871

网　　　址：http://www.pup.cn　新浪官方微博：@北京大学出版社

电 子 信 箱：pup_6@163.com

电　　　话：邮购部 62752015　发行部 62750672　编辑部 62750667　出版部 62754962

印　刷　者：

经　销　者：新华书店

787 毫米×1092 毫米　16 开本　20.75 印张　483 千字

2014 年 3 月第 1 版　2014 年 3 月第 1 次印刷

定　　　价：42.00 元

前　　言

当前，无线通信在世界各地都得到了快速发展，并已成为全球通信和 IT 界共同关注的热门领域，无线通信涉及的技术内容非常广泛，尤其随着物联网建设的全面铺开，无线通信的基础技术不断完善和发展变化，各种不同类型的无线通信系统和无线通信新技术更是不断涌现。因此，在选材上，本书主要关心无线通信领域涉及的基本理论、技术和原理，以及移动通信网络和无线通信的最新技术；内容上力求突出重点、丰富新颖、深入浅出、通俗易懂。本书适合高等院校通信工程、电子工程、信息工程、新兴的物联网工程及其他相关专业本科高年级学生作为教材使用，也可供通信工程技术人员和科研人员作为参考书使用。

全书共分 8 章。第 1 章绪论，主要介绍现代无线通信系统的组成、分类和特点、无线通信方式、无线通信的主要技术指标和几种典型的无线通信系统等。第 2 章无线信道传输，主要介绍电磁波的基本知识和无线信道的传输特性、典型的随参信道、随参信道的特点及传输特性、随参信道对信号传输的影响、电磁波在随参信道中传输产生的各种衰落及几种抗衰落技术。第 3 章无线通信编码技术，主要介绍信源编码、信道编码、线性分组码、循环码、卷积码、Turbo 码等。第 4 章无线调制技术，主要介绍调制的概念、模拟调制(包括幅度调制和角度调制 AM、FM、PM 等)的原理、数字调制技术(包括 ASK、FSK、PSK、DPSK、QPSK、MSK、GMSK、OFDM、扩频等)的调制解调原理。第 5 章多址接入技术，主要介绍频分多址(FDMA)、时分多址(TDMA)、空分多址(SDMA)、码分多址(CDMA)、随机多址技术及其他几种多址接入技术。第 6 章无线通信协议，主要介绍不同无线通信网络的无线通信协议标准，包括无线城域网通信协议、无线局域网通信协议、无线个域网通信协议、无线传感器网络中典型的 MAC 层协议和路由协议、无线 Mesh 网络中典型的 MAC 协议和路由协议。第 7 章蜂窝移动通信，主要介绍移动通信的概念和特点、GSM 移动通信系统、CDMA 移动通信系统、第三代移动通信系统及其四大技术标准(WCDMA、CDMA2000、TD-SCDMA 和 WiMAX)，以及第四代移动通信的主要技术及其未来发展。第 8 章无线通信新技术，对近几年无线通信领域普遍关心的新技术进行了简单介绍，其中包括软件无线电技术、认知无线电技术、智能天线技术、超宽带无线技术等，以便于读者跟踪无线通信技术的最新发展。

本书由许晓丽和赵明涛担任主编，董光辉和刘靖宇担任副主编。第 1、2、4 章由许晓丽编写，第 3、5 章由赵明涛编写，第 7 章和全书习题及习题答案由董光辉编写，第 6、8 章由刘靖宇编写。全书由许晓丽修改定稿。本书在编写过程中，得到了北京大学出版社的大力支持，在此一并表示感谢！

鉴于作者水平有限，书中难免存在不妥之处，恳请读者批评指正。

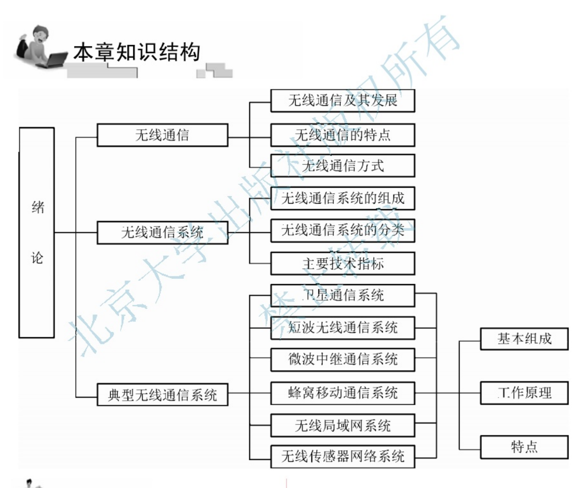
编　者

2013 年 11 月

目　录

第 1 章 绪 论

本章知识结构



本章教学目的与要求

- 掌握无线通信的基本概念
- 掌握无线通信系统的基本组成及各部分的功能
- 掌握无线通信系统的分类和通信方式
- 掌握无线通信的主要技术指标
- 了解无线通信的发展、特点以及无线通信的应用
- 了解几种典型的无线通信系统及其组成和特点



引言

随着社会的发展，人们对通信的需求日益迫切，对通信的要求也越来越高，无线通信让人们摆脱有线网络的束缚，以实现任何人(Whoever)在任何时间(Whenever)、任何地点(Wherever)与任何人(Whomever)进行任何种类(Whatever)的信息交换(即 5 个“W”)为目标。因此作为有线通信的补充和发展，无线通信自 21 世纪初以来得到了迅猛的发展，并且已经成为当今社会不可缺少的信息交流手段。本章将主要介绍无线通信及其系统的基本组成、发展、特点，无线通信系统的分类、通信方式和几种典型的无线通信系统。

【案例 1.1】无线通信的诞生

无线通信始于 19 世纪后期，英国的麦克斯韦是人类历史上第一个预言电磁波存在的人，他于 1864 年提出了电磁场的动力学理论，而后他又于 1873 年出版了科学名著《电磁理论》，系统、全面、完美地阐述了电、磁、光相统一的电磁波理论，这一理论成为经典物理学的重要支柱之一，并为无线通信奠定了坚实的理论基础。德国的赫兹是第一个通过实验证实电磁波存在的人。赫兹按照麦克斯韦的理论，精心设计了“振荡偶极子”(最原始的无线电发射机)和“共振偶极子”(最原始的无线电接收机)。经过几年的努力，1887 年，赫兹通过实验终于成功地探测到电磁波的存在——在离振荡偶极子 3 米远的地方，观察到了在共振偶极子的两个小铜球的隙缝中跳跃的电火花，这种微弱的电火花证实了麦克斯韦所预言的电磁波的存在。有了电磁波，才使得无线通信成为可能。

意大利的马可尼是实现无线通信应用的创始人，1897 年马可尼的横跨布里斯托尔海峡无线电通信试验取得成功，这标志着无线通信的诞生。1897 年，马可尼在伦敦成立了由英国政府代办的马可尼无线电报公司，他亲自兼任董事长。有了资金的支持，马可尼开始把无线电通信向商业化发展。1898 年，英国举行游艇赛，终点是距海岸 20 英里的海上。《都柏林快报》特聘马可尼用无线电传递消息，游艇一到终点，他便通过无线电波，向岸上的观众及时通报了比赛的结果，使观众们欣喜若狂，这也是无线电通信的第一次实际应用。1899 年 10 月，马可尼在怀特岛上播发了第一封收费电报，电码飞越英国和法国之间的英吉利海峡，通信距离达 45km，这标志着无线通信成功地实现了商用。图 1.1 是马可尼发明的第一台无线电收发装置。事实上，无线电的发明者并不只是马可尼一人。俄国的波波夫，差不多与马可尼同时，也独立地完成了无线电报通信的发明，并制造了第一台无线电接收机，图 1.2 是波波夫发明的第一台无线电接收机。无线通信还和许多发明家的名字联系在一起，对于无线通信技术的贡献，集中了那个时代很多人的智慧。在今天的信息社会中，他们发明的意义随着时间的流逝而愈显珍贵。可以说，如今的无线通信，使得神话中的“千里眼”和“顺风耳”都成为了现实。无线通信业务和应用的种类日益繁多，现代的无线通信发展到了什么样子呢？又有哪些无线通信新技术？无线通信未来的发展又如何呢？通过本章的学习将得到答案。

【案例 1.2】物联网——泛在的通信世界

我们正站在一个新的通信时代的边缘，物联网使我们在信息与通信技术的世界里获得一个新的沟通维度，信息与通信技术的目标已经从满足人与人之间的沟通，发展到实现人

与物、物与物之间的连接，世界上所有的物品，小到纽扣、手表，大到汽车、公路设施、楼房等都可以通过互联网进行通信，并实现智能化识别和管理，万物的连接就形成了物联网，目前，物联网已在智能家居、智能医疗、智慧城市、智能交通、智能校园等领域得到了广泛应用。

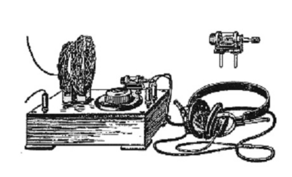


图 1.1　马可尼发明的第一台无线电收发装置

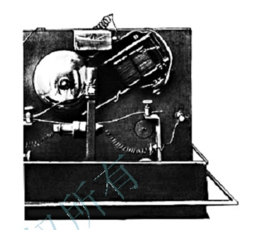


图 1.2　波波夫发明的第一台无线电接收机

物联网通过 RFID 技术、无线传感、无线定位等技术自动识别采集和感知获取物品的标识信息，物品自身的属性信息和周边环境信息，借助各种信息传输技术将物品信息聚合到统一的信息网络之中。葡萄什么时候成熟、什么时候可以采摘，不需要到地里，通过传感器，坐在计算机前便可掌握；下班前轻轻按下手机的按键，家里已设定的空调便会在进门前 20 分钟自动启动，一进家门就可享受凉爽；在手机里安装“掌上公交”客户端，可随时查询城市各条线路公交车的实时位置、何时到站离站等信息；在“未来商店”试衣服成为一种新奇的享受……。这些看起来很炫的生活，因为物联网(图 1.3)——这个泛在的通信世界的出现而正在我们身边悄悄实现。

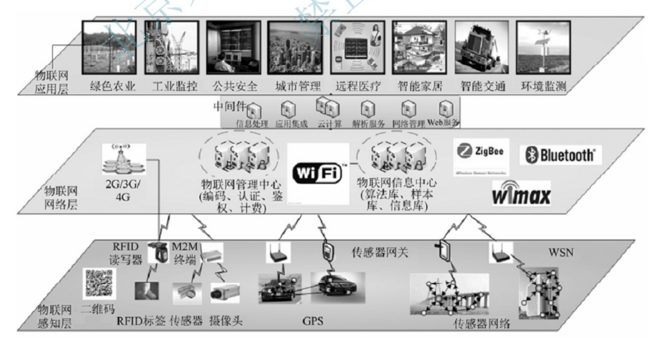


图 1.3　物联网——泛在的通信

1.1 无 线 通 信

1.1.1 无线通信及其发展

无线通信是通过电磁波在自由空间传播以实现信息传输为目的的通信。无线通信的通信双方至少有一方以无线方式进行信息的交换和传输。无线通信可用来传输电报、电话、传真、图像数据和广播电视等通信业务。与有线通信相比，无线通信无须架设传输线路、不受通信距离限制、机动性能好、建立迅速。

自 19 世纪后期无线通信诞生后，20 世纪的第一个十年期间，无线电通信并未获得普及应用，多半局限于在船舶等移动物体上使用，而且无线电报的通信距离是很有限的，通信的可靠性也较差。电磁波在向空间传输过程中，随着传输距离的增加，信号强度急剧减弱。显然，要扩大无线电报通信距离，就要增大发射机的发射功率和提高接收机的灵敏度。第一代电子器件的发明改变了有线通信独占通信领域的局面，1904 年，英国弗莱明发明了真空二极管。真空二极管应用于无线电接收装置后，接收机的灵敏度和可靠性大为提高。1906 年到 1907 年，美国德福雷斯特发明了真空三极管，此后经过不断改进，真空三极管在无线电通信中得到了广泛应用。电子管的发明，使得无线电话、无线广播真正可以获得实用。1906 年，美国发明家森登成功地研究出人类历史上第一次不用导线而用电磁波传送语言和音乐的无线电广播，被称为“载着声音飞翔的电波”，但声音作用于送话器所变换成的音频电信号十分微弱，这制约了通话的距离和声音传输的质量。只有当电子管进入实用阶段后，人们才可以借助电子管对送话器输出的微弱音频电信号进行放大，然后对高频无线电波进行调制，无线电话才真正具有了生命力。

泰坦尼克号惨案进一步推动了无线通信应用的普及。1912 年 4 月，英国新建的当时世界上最豪华的邮轮泰坦尼克号，在其开往美国的处女航行中因冲撞冰山而沉没，死难者达 1500 多人。泰坦尼克号配备了较完善的无线电报装置，在它发出 SOS 国际无线电呼救信号时，在距它 40km 的洋面上恰好有一艘货轮经过，但是这艘货轮却未曾安装无线电报设备。等到距离泰坦尼克号 100km 的卡尔巴夏号轮船接到信号赶赴出事地点时，只抢救出 700 多人，在死难的一千多人中，有不少是因为久久没有得到救助而经受不住饥寒袭击，惨死于漂浮在冰海的救生艇上。全世界从这一惨剧中吸取了教训，此后各国法律强行规定：凡是具有一定规模的船只必须配备无线电装置。

1916 年，美国遇到暴风雪袭击，有线电报线路中断，从此开始，无线电报逐渐应用于火车调度。各种类型的电台也如雨后春笋般地蓬勃发展起来，并被广泛地用于传递商品行情、军事情报、气象消息和新闻等。后来，无线电通信逐渐又被用于战争。在第一次和第二次世界大战中，无线通信都发挥了很大的威力，以致有人将第二次世界大战称之为“无线电战争”。

1920 年美国无线电专家康拉德在匹兹堡建立了世界上第一家商业无线电广播电台，从此广播事业在世界各地蓬勃发展，收音机成为人们了解时事新闻的方便途径。1924 年第一条短波通信线路在瑙恩和布宜诺斯艾利斯之间建立。1933 年法国人克拉维尔建立了英法之间的第一个商用微波无线电线路，推动了无线通信技术的进一步发展。

20 世纪 30 年代，无线通信系统的规模应用始于美国，这些早期的无线通信采用调幅(AM)制式技术，由于 AM 技术的特点，早期无线通信系统的抗幅度干扰能力很弱。随着后来的频率调制(FM)技术的问世及其快速发展和应用，早期的无线移动通信系统抗幅度干扰的能力大大增强，因此，调频技术成为无线通信的主要调制技术。

1968 年贝尔实验室提出的蜂窝电话技术概念，是无线通信技术发展历史上的一项重大突破。采用蜂窝概念可以设计任意大容量的无线通信系统，从而解决了无线通信技术向公众服务应用发展的关键问题。1979 年，第一代模拟蜂窝系统 NTT 首先在日本投入市场应用。1983 年，AMPS 系统在美国投入商用。1985 年，ETACS 在欧洲投入商用。第一代蜂窝系统均采用频分多址的模拟射频传输组网技术，主要用于话务通信。由于第一代移动通信系统存在系统容量有限、业务种类单一且质量差、抗干扰能力弱、不能传送数字信息等缺点，所以第一代移动通信系统已经停用。

20 世纪 80 年代以后，随着超大规模集成电路技术和数字信号处理技术的进步、蜂窝技术和无线接入技术的飞速发展，无线通信技术及其应用进入快速发展阶段。第二代移动通信系统主要采用时分多址(TDMA)和码分多址(CDMA)数字接入技术，支持语音通话、收发电子邮件等低速率数据业务，从 2G 开始，无线通信步入了纯数字时代，所有的标准都以商业利益为宗旨。在这个发展过程中，先是 GSM 提供了短数据业务(SMS)，而后发展到称为 2.5G 的 GPRS。与 GSM 对应但发展稍晚的另一个标准——CDMA 也经历着相似的发展演进过程。目前，世界上还有一部分 2G 无线移动通信系统仍在运营之中，2G 标准包括 GSM、iDEN、USDC、IS-95、PDC、CSD、PHS、GPRS、HSCSD 和 WiDEN。

为了满足不断增长的网络容量和高速数据传输的需求，无线通信又发展到了第三代宽带多媒体移动通信(3G)阶段，3G 系统基本上是 2G 的线性扩展，主要是以 CDMA 技术为核心，提供高速的数据传输能力(最高达到 2Mbps)，提供移动 Internet 和多媒体业务、视频会议、可视电话等。常见的 3G 技术标准有 WCDMA、CDMA2000 和 TD-SCDMA，在我国，3G 已于 2008 年投入使用。

目前超 3G(B3G)技术和 4G 技术也已经开始研究，3G 数据速率可以达到 2Mbps，B3G 数据速率则可以达到 1Gbps 以上。未来的第四代移动通信(4G)将采用正交频分复用技术、智能天线技术和切换定位等新技术，集成了不同模式的无线通信用户，可以自由地从一个标准漫游到另一标准，具有更高的数据率、高频谱利用率、低发射功率和灵活的业务支撑力能力，其支持业务更广，如宽带移动和无缝业务，宽带移动包括基本的语音、视频和移动 Internet，分组交换的语音业务和数据业务，远程通信、虚拟现实等；无缝业务是要使现存的网络与新建网络很好地互联、融合，并保证终端漫游中的不间断性、跨多个网络时业务提供的透明性。

随着无线通信技术的飞速发展，现代无线通信的应用范围越来越广泛，如无线电广播、电视、移动通信、航空航海、舰船、导航、雷达、天文学、目标的追踪和定位、卫星通信、微波中继、宽带无线接入、宽带无线检测、射频识别(RFID)、无线传感器网络、物联网等。

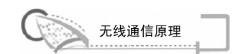


图 1.4 为现代无线通信网的组织结构示例，它应用现代无线通信技术实现了各种网络的接入和互联。从无线局域网、无线个域网、无线城域网到无线广域网，从移动 Ad hoc 网络到无线传感网络、无线 Mesh 网络、从 WiFi 到 WiMedia、WiMAX，从 IEEE 802.11、IEEE 802.15、IEEE 802.16 到 IEEE 802.20，从 GSM、GPRS、CDMA 到 3G、超 3G、4G 等，多种网络融合在一起，使之相互取长补短，发挥每一种网络的长处，从而逐步实现与完善符合未来个人通信需求的综合性通信网络。

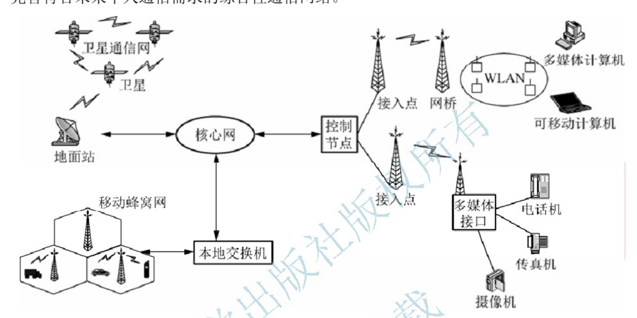


图 1.4　现代无线通信网的组织结构示例

1.1.2　无线通信的特点

无线通信以电磁波辐射的形式传播信息，传输信道为自由空间，省去了铺设线缆的费用，使用方便，接入灵活，通信距离远，不受地理条件限制等，然而无线通信是开放的信道，信道中的干扰无法得到有效控制，且无线通信终端往往所处位置各不相同，故无线通信较复杂。

1. 复杂的无线传播环境

无线通信具有复杂的无线传播环境。有线通信使用特性稳定的传输媒介，传输环境是稳定的和可预测的，而无线通信使用无线信道作为传输媒介，传输环境复杂多变。在发射机到达接收机的传播路径上，很少出现简单的视距传播情况，多数情况下，电磁波在传播过程中会受到城市高大建筑物的阻挡，受到许多地物的反射、绕射或散射的影响，导致发射机发出的同一信号可能经过多条路径到达接收机，这就是多径传播现象。多径传播使电磁波的传播衰减增大，还会产生严重影响通信效果的多径衰落。

2. 系统复杂

无线通信系统的固有特性之一是用户的可移动性，这一特性对系统设计有着重要的影

响，人们熟知的蜂窝系统就是典型的例子。一方面，移动增加了无线信道的复杂性，导致无线信道是一个时变的多径信道；另一方面，蜂窝系统对移动用户的管理也是一个比较复杂的过程，在任何时候系统都需要确定用户的位置，采用频率和功率控制、地址登记等跟踪技术来跟踪用户。由于无线通信用户移动性的特点，对计费及结算提出了更高要求，随着通信技术的不断发展，无线通信业务日趋完善并呈现多样化趋势，多种业务类型就要求多种计费方式，尤其是数据业务，需按时长计费、按流量计费、按内容计费等，这就更增加了系统的复杂性。此外，无线终端长期处于不固定状态，外界的影响很难预料，尘土、振动、日晒雨淋的情况时常遇到，这就要求移动终端性能稳定可靠、小型便于携带、低功耗及能耐高温和低温等，这也给无线终端的设计和制造带来较大的困难。

3. 有限的频谱资源

有限的频率资源决定了有限的信道数目，这与日益增长的用户数量形成了一对矛盾。无线电频谱是一种资源，这种资源具有以下主要特点：一是有限性，其空间、时间和频率三维要素可以重复使用，但是在一定条件下对某一频段和频率的利用又是有限的；二是非耗竭性，频率资源不同于土地、水、矿产等一类再生或非再生资源，不利用是一种浪费，使用不当也是一种浪费，甚至会造成危害；三是固有性，它的传播不受行政区域的限制，既无省界也无国界。由于无线电频谱的这些特点，无线频谱的使用是通过国际协议进行管制的，在中国是由国家无线电管理部门进行管理的。

4. 多种干扰并存

在无线通信系统中，空间传播的电磁波除了有用信号外，还存在大量的无线电干扰，无线电干扰是指对有用信号造成损坏的无用信号或电磁干扰，它会引起无线电通信系统的接收性能下降、误码或信息丢失。无线电干扰通常按干扰的性质分为人为干扰和自然干扰。自然干扰来源于太阳干扰、宇宙干扰和天电干扰等，是不可控制的。人为干扰又可细分为无线电设备干扰和非无线电设备干扰。非无线电设备干扰包括工业、科研、医疗等电器设备干扰和电力线干扰等，为防止其对无线电设备产生有害干扰，国家标准中已规定了其使用频率和辐射允许值的大小。无线电设备干扰是无线电通信的主要干扰，它主要包括互调干扰、邻道干扰和同频干扰。互调干扰是指两个或多个干扰信号同时加到接收机时，由于非线性的作用，这两个干扰的组合频率有时会恰好等于或接近有用信号频率而顺利通过接收机，如偶尔收音机里的杂音和串台就是由于这个原因造成的。邻道干扰是指相邻的或邻近频道的信号相互干扰，一般是由于一个强信号串入弱信号中干扰弱信号而造成的影响。例如，我们知道调频信号的频谱是很宽的，理论上说，调频信号含有无穷多个边频分量，当其中某些边频分量落入邻道接收机的通带内，就会造成邻道干扰。为解决这个问题，在无线通信设备中常常采用自动功率控制技术，对强功率信号加以控制，同时使用滤波器以减少邻频的泄露。同频干扰是无用信号的载频与有用信号的载频相同，并对接收同频有用信号的接收机造成的干扰，它是电磁波在无线环境下传输时特有的干扰，由无线资源重复利用所造成的。

1.1.3 无线通信方式

通信方式是指通信双方或多方之间的工作形式和信号传输方式。根据不同的标准，无线通信方式也有不同的分类。

(1) 按通信对象数量的不同，通信方式可分为点与点之间的通信(两个对象之间的通信)、一点与多点之间的通信(一个对象和多个对象之间的通信)和多点与多点之间的通信 3 种(多个对象和多个对象之间的通信)，这 3 种通信方式的示意图如图 1.5 所示。

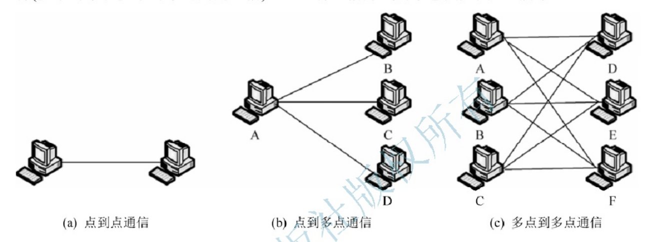
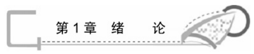


图 1.5 通信对象数量不同的通信方式示意图

(2) 按信号传输方向与传输时间的不同，任意两点间的通信方式可分为单工、半双工和双工通信方式，3 种通信方式示意图如图 1.6 所示。单工通信是指在任何一个时刻，信号只能从甲方向乙方单向传输，甲方只能发信，乙方只能收信，如广播电台与收音机、电视台与电视机的通信、遥控玩具、航模、寻呼等均属于单工通信。半双工通信是指在任何一个时刻，信号只能单向传输，或从甲方向乙方，或从乙方向甲方，每一方都不能同时收、发信息，如对讲机、收发报机之间的通信。双工通信是指在任何一个时刻，信号能够双向传输，每一方都能同时进行收信与发信工作，如普通电话、手机之间的通信。

(3) 按信号传输顺序的不同(主要指数字通信)，通信方式可分为串行通信与并行通信。串行通信是指将表示一定信息的数字信号序列按信号变化的时间顺序一位接一位地从信源经过信道传输到信宿。并行通信是指将表示一定信息的数字信号序列按码元数分成 n 路(通常 n 为一个字长，如 8 路、16 路、32 路等)，同时在 n 路并行信道中传输，信源一次可以将 n 位数据传送到信宿。如在传输数字信号 10011010 时，并行方式则将该序列的 8 位码用 8 路信道同时传输，如图 1.7 所示。

(4) 按同步方式的不同，通信方式可分为同步通信和异步通信。异步通信以字符为通信单位，同步信息由硬件加在每一个字符的数据帧上。与异步通信不同，同步通信不是对每个字符单独同步，而是以数据块为传输单位并对其进行同步。每个数据块的头部和尾部都要附加一个特殊的字符或比特序列，以标志数据块的开始与结束，这里数据块是指由一批字符或二进制符号序列组成的数据。



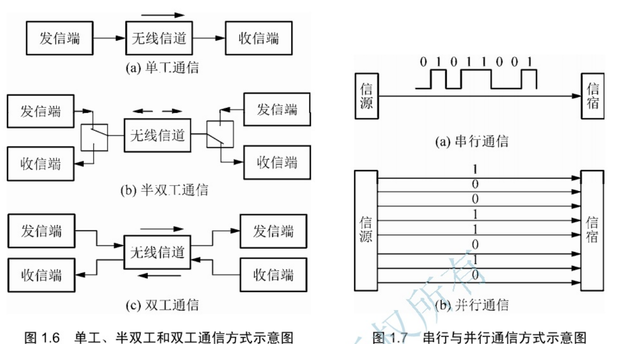


图 1.6 单工、半双工和双工通信方式示意图

图 1.7 串行与并行通信方式示意图

1.2 无线通信系统

1.2.1 无线通信系统的组成

不同的无线通信系统，虽然它们具体的设备组成和复杂度差异较大，但基本组成都是一样的，图 1.8 给出了无线通信系统的基本组成框图，它包括信源、发送设备、无线信道、噪声与干扰、接收设备、信宿这六大基本组成部分。

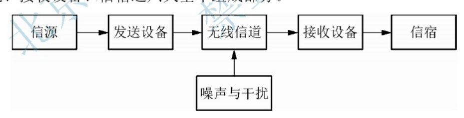


图 1.8 无线通信系统的基本组成

信源是发出信息的基本设备，它的主要作用是将待发送的原始信息变换为电信号，这种电信号也称为基带信号。例如，话筒将声音变为电信号，还有如摄像机、电传机和计算机等设备都可以看作信源。

发送设备是将信源产生的电信号转换成适合在无线信道中传输的电磁波信号，并将此电磁波信号送入无线传播信道，从而将信源和无线信道匹配起来。发送设备一般包括两方面的功能，即调制和放大。放大包括电压放大和功率放大，放大的主要目的是提高发送信号的功率。在需要频谱搬移的场合，调制是最常见的变换方式。调制将低频信号加载到高频载波中，从而实现信号的远距离、多路、低损耗的快速传输。调制可以通过使高频载波信号随基带信号的变化而改变载波的幅度、频率或相位来实现。调制方式可分为模拟调制

与数字调制两大类，第一代无线通信系统采用模拟调制，目前的无线通信系统都采用数字调制。常用的数字调制方式有 ASK、FSK、PSK、MSK、GMSK、QPSK、8PSK、16QAM、64QAM 等。本书后面的章节将会详细介绍这些数字调制方式。对数字无线通信系统来说，发送设备还包括信源编码和信道编码。信源编码将来自信源的连续消息变换为数字信号，并对其进行适当的压缩处理以提高传输效率。信道编码使数字信号与无线传输信道相匹配，通过在被传输数据中引入冗余来避免数据在传输过程中出现误码，目的是提高传输的可靠性和有效性。用于检测错误的信道编码称为检错编码，既可检错又可纠错的信道编码称为纠错编码。常见的信道编码方式有分组码、卷积码、Turbo 码、循环码等。

无线信道是电磁波传输的通道，对于无线通信来说，无线信道主要是指自由空间，也包括水等。对于电磁波而言，它在发送端与接收端之间的无线信道中传输时，并没有一个有形的连接，其传播路径也往往不止一条，因此电磁波在传输过程中必然会受到多种干扰的影响而产生各种衰落，从而造成系统通信质量的下降。

噪声与干扰是无线通信系统中各种设备及信道中所固有的，它不是人为加入的设备，并且是人们所不希望的。对于无线通信，信道中的噪声和干扰对信号传输的影响较大，是不可忽略的，为了分析方便，它被看成是各处噪声的集中表现而抽象加入到无线信道中的一部分。

接收设备的功能与发送设备的功能相反，主要是接收自由空间中传输过来的电磁波，从带有干扰的接收信号中正确还原出相应的原始基带信号。接收设备具体包括解调、译码、解码等功能。此外，在发送设备和接收设备中需要安装天线来完成电磁波的发送和接收。

信宿是信息传输的归宿点，其作用是将还原的原始基带信号转换成相应的原始信息。

1.2.2 无线通信系统的分类

无线通信系统根据不同的原则可以有不同的分类。

根据传输信号的形式分类：电话通信系统、电报通信系统、广播通信系统、电视通信系统等。

根据无线应用分类：移动通信系统、无线接入通信系统、微波通信系统、卫星通信系统等。

根据无线终端(用户)工作状态分类：固定无线通信系统和移动无线通信系统。

根据工作频段分类：长波通信、中波通信、短波通信、超短波通信、远红外无线通信、微波通信和卫星通信等。工作频率主要指发送与接收的射频(RF)频率。射频实际上就是广义的“高频”，它是指适合无线电发射和传播的频率，无线通信未来发展的方向之一就是开辟更高的频段。表 1-1 给出了无线电波频段的划分。

表 1-1　无线电波频段的划分

波段名称	频段名称	频率范围	波长范围	主要用途
超长波	甚低频 VLF	3～30kHz	10～100km	海岸潜艇通信、远距离陆地通信、超远距离导航等
长波	低频 LF	30～300kHz	1～10km	越洋通信、中距离通信、地下岩层通信、远距离导航等

续表

波段名称	频段名称	频率范围	波长范围	主要用途
中波	中频 MF	300kHz～3MHz	0.1～1km	船用通信、业余无线电通信、移动通信、调幅广播等
短波	高频 HF	3～30MHz	10～100m	远距离短波通信、国际定点通信
超短波	甚高频 VHF	30～300MHz	1～10m	电离层散射、流星余迹通信、电视、调频广播、空中管制、导航等
分米波、微波	特高频 UHF	300MHz～3GHz	10～100cm	小容量微波中继通信、卫星通信、空间通信、雷达等
厘米波	超高频 SHF	3～3GHz	1～10cm	大容量微波中继通信、卫星通信、空间通信、雷达等
毫米波	极高频 EHF	30～300GHz	1～10mm	雷达、微波接力、射电天文学、波导通信等

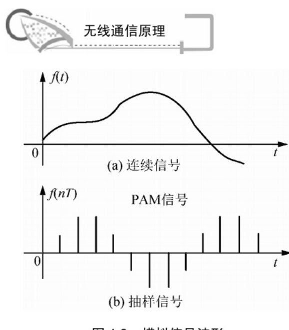

小知识

射频识别技术是物联网中的关键技术之一，射频识别系统(RFID)需要产生并辐射电磁波，但 RFID 系统要顾及其他无线电服务，不能对其他无线电服务造成干扰，因此 RFID 系统通常使用为工业、科学和医疗特别保留的 ISM 频段。通常使用的 ISM 频段为 6.78MHz、13.56MHz、27.125MHz、40.68MHz、433.92MHz、869.0MHz、915.0MHz、2.45GHz、5.8GHz 及 24.125GHz 等，RFID 常采用上述某些 ISM 频段，除此之外，RFID 也采用 0～135kHz 的频率。

根据调制方式分类：调幅、调频、调相及混合调制等无线通信系统。

根据无线信道传输的信号参量取值的不同分类：模拟无线通信系统和数字无线通信系统。凡信号参量的取值是连续的或取无穷多个值的，且直接与消息相对应的信号称为模拟信号，如电话机送出的语音信号、电视摄像机输出的图像信号等。模拟信号有时也称连续信号(见图 1.9(a))，这个连续是指信号的某一参量可以连续变化，或者说在某一取值范围内可以取无穷多个值，而不一定在时间上也连续，如图 1.9(b)所示的抽样信号。凡信号参量只能取有限个值，并且常常不直接与消息相对应的信号称为数字信号，如电报信号、计算机输入/输出信号、PCM 信号等。数字信号有时也称离散信号，其二进制波形如图 1.10(a)所示。这里的离散是指信号的某一参量是离散变化的，而不一定是在时间上离散变化，如图 1.10(b)所示的 2PSK 信号。

根据信号的复用方式分类：频分复用、时分复用和码分复用无线通信。频分复用是用频谱搬移的方法使不同信号占据不同的频率范围；时分复用是用脉冲调制的方法使不同信号占据不同的时间区间；码分复用是用正交的脉冲序列分别携带不同信号。传统的模拟通信中都采用频分复用，随着数字通信的发展，时分复用通信系统的应用越来越广泛，码分复用主要用于空间的扩频通信中。



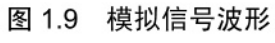


图 1.9 模拟信号波形

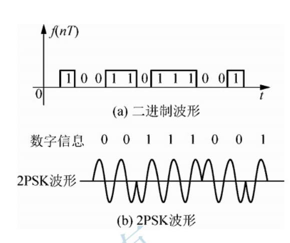


图 1.10 数字信号波形

1.2.3 无线通信系统的主要技术指标

无论是哪种类型的通信，其主要任务都是实现信息快速、准确的传输，因此衡量一个通信系统好坏的主要性能指标就是通信的有效性和可靠性。有效性是指通信系统中信息传输的快慢问题；而可靠性是指信息传输的准确性问题。有效性指标与可靠性指标在不同的通信系统中有着不同的表述方法，在无线通信系统中，有效性和可靠性的主要技术指标包括信道带宽、传输速率、频带利用率、差错率等。

1. 无线信道带宽

带宽，即频带宽度，一般是指波长、频率或能量带的范围，用 B 表示，单位是 Hz。不同的带宽，其含义是不同的，因此实际用到带宽时，往往需要说明是哪种带宽。通常带宽分为信道带宽和信号带宽，信道带宽是一个信道能够传送电磁波的有效频率范围；信号带宽是指信号所占据的频率范围，由信号的特点决定。对于无线通信来说，无线信道传输的信号都是具有一定频带宽度的信号，通常信号带宽是小于信道带宽的，香农理论中的信道容量定理决定了传输信号的最大带宽。无线信道带宽除了与信道的特性有关外，还与国际频段管理组织的频段划分与通信体制有关。



小知识

无线频率的使用是由政府统一管理的，无线通信系统的工作频段是指政府无线电管理部门划分给该项业务的工作频率范围。在政府划分的工作频段上，运营商需要按照无线通信系统的单载波信道带宽进一步将可用频段细分。

中国移动通信公司的GSM网络的单载波信道带宽为200kHz，CDMA的单载波信道带宽为1.25MHz，TD-SCDMA 的单载波信道带宽为 1.6MHz，WCDMA 的单载波信道带宽为 5MHz。可用频段划分成一系列单载波无线信道之后，运营商要根据自己的网络建设方案将这些无线信道编组，并确定合适的频率复用方案。

2. 传输速率

1) 信息传输速率

信息传输速率通常也称为传信率或比特率等，它表示无线信道中单位时间内传输信息的比特数，单位是比特/秒(bps)。

例如，GSM 系统信道速率为 270.833Kbps，超宽带(UWB)无线通信系统的信道速率可达几百兆比特/秒。

2) 码元传输速率

码元传输速率通常也称为码元速率、数码率、传码率或波特率等，它表示无线信道中单位时间内传输码元的数目，单位是波特(Baud)。通常波特率是小于等于比特率的，二进制数据传输时，码元传输速率与信息传输速率相等。

3) 最大传输速率

最大传输速率：每个信道传输数据的速率有一个上限即为信道的最大传输速率，也就是信道容量，它是根据香农理论确定的。

3. 频带利用率

频带利用率是指单位频带内的传输速率，它是更能准确衡量无线通信系统传输有效性的技术指标。信息传输需要一定的频带，传输系统所占的频带越宽，传输数据信息的能力就越强，因此比较不同信息传输系统的有效性时，只考虑它们的传输速率是不够的，还应看在这样的传输速率下所占的信道的频带宽度。通常频带利用率有两种表示方法：一是码元传输速率意义下的频带利用率，另一个是信息传输速率意义下的频带利用率，即

$$\eta = \frac{\text{码元传输速率}}{\text{频带宽度}} \tag{1-1}$$

$$\eta = \frac{\text{信息传输速率}}{\text{频带宽度}} \tag{1-2}$$

4. 差错率

差错率是衡量无线通信系统可靠性的重要指标，常用误码率和误信率来描述。

误码率也称码元差错率，它表示码元在传输系统中被传错的概率，即发生差错的码元数占传输总码元数的比例

$$P_{\mathrm{e}} = \frac{\text{发生错误的码元数}}{\text{传输总码元数}} \tag{1-3}$$

误信率也称信息差错率，它表示发生差错的比特数占传输总比特数的比例，即

$$P_{\mathrm{b}} = \frac{\text{发生错误的比特数}}{\text{传输总比特数}} \tag{1-4}$$

5. 其他指标

1) 吞吐量

吞吐量是信道在单位时间内成功传输信息的数量，其单位为 bps。由于传输过程中出错或丢失数据造成重传的信息量，不计在成功传输的信息量之内。

例如，某信道 10min 内成功传输了 8.4Mbit 的数据，那么它的吞吐量就是 8.4Mbit/600s = 14Kbps 。

2) 延迟

延迟是指从发送端发送第一位数据开始到接收端成功地接收到最后一位数据为止所经历的时间。它主要分为传输延迟和传播延迟两种，传输延迟与数据传输速率和发送机、接收机以及中继和交换设备的处理速度有关，传播延迟与传播距离有关。

3) 抖动

延迟不是固定不变的，它的实时变化称为抖动。抖动往往与机器处理能力、信道拥挤程度等有关。有的应用对延迟敏感，如电话；有的应用对抖动敏感，如实时图像传输。

1.3 无线通信系统实例

1.3.1 卫星通信系统

卫星通信是指利用人造地球卫星作为中继站转发无线电信号而实现两个或多个地面站之间的通信。卫星通信是现代无线通信的重要手段。目前全世界有 200 多个国家和地区的约 200 多颗同时在地球静止轨道上运行的通信卫星，它们提供了 80%的洲际通信和 100%的国际电视转播以及国内或区域内的通信和电视广播业务。

由于卫星处于外层空间，即地球在电离层之外，地面上发射的电磁波必须能穿透电离层才能到达卫星，同样，从卫星到地面上的电磁波也必须穿透电离层。电磁波穿过大气层时，必然要受到大气层的吸收而使信号的能量损耗和衰减。人们经过大量的分析和测试，总结出电磁波穿过大气层时的损耗与使用的频率有关，并且只有在微波波段时，大气层对电磁波的吸收最小，电磁波能量损耗最小，所以微波波段信号穿过大气层的能力最强，因此微波波段又被称为无线电窗口。目前大多数卫星通信系统都工作在微波频段 0.3～30GHz，具体频段有 UHF 波段(0.3～1.0GHz)、L 波段(1.0～2.0GHz)、S 波段(2.0～4.0GHz)、C 波段(4.0～8.0GHz)、X 波段(8.0～12.0GHz)、Ku 波段(12～18GHz)、K 波段(18～27GHz)、Ka 波段(27～40GHz)。由于 C 波段的频段较宽，又便于利用成熟的微波中继通信技术，且天线尺寸也较小，因此，卫星通信最常采用的是 C 波段。

卫星通信系统根据通信卫星与地面之间的位置关系可分为静止通信卫星、移动通信卫星、中地球轨道卫星和低地球轨道卫星。

静止通信卫星：静止通信卫星也称为同步通信卫星，它是轨道位于赤道平面上的卫星，它离地面高度为 35780km，相对于地面位置是静止的。采用 3 个相差 120° 的静止通信卫星就可以覆盖地球的绝大部分地域(两极盲区除外)，如图 1.11 所示。

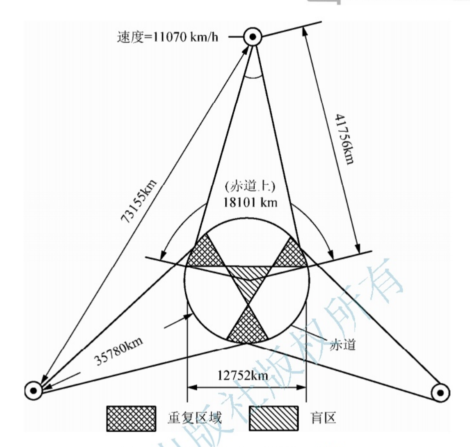


图 1.11　静止通信卫星覆盖地球示意图

移动通信卫星：椭圆形轨道，相对于地面位置是移动的。

中地球轨道卫星：距地面 5000～20000km，卫星运行周期 4～12h，相对于地面位置是移动的。

低地球轨道卫星：距地面 500～5000km，卫星运行周期 2～4h，相对于地面位置是移动的。卫星通信系统由通信卫星、地球站、上行线路及下行线路组成。通信卫星是实现通信目的的人造地球卫星，上行线路和下行线路是地球站至通信卫星及通信卫星至地球站的无线电传播路径，并且上行线路和下行线路的工作频率是不同的，通常上行频率高于下行频率。通信设备主要集中于地球站和通信卫星中。卫星通信系统组成示意图如图 1.12 所示。

卫星通信系统线路组成框图如图 1.13 所示。为了进行双向通信，每一个地面站均应包括发射系统和接收系统。由于收、发系统一般是共用一副天线，因此需要使用双工器来将收、发信号分开。地面站收、发系统的终端，通常都是与长途电信局或微波线路连接。地面站的规模大小则由通信系统的用途而定。在卫星通信系统中，各地面站发射的信号都是经过卫星转发给对方地面站的，卫星转发器的作用是接收地面站发来的信号，经变频、放大后，再转发给其他地面站。因此，除了要保证在卫星上配置转发无线电信号的天线及通信设备外，还要有保证完成通信任务的其他设备。一般来说，一个通信卫星主要由天线、接收设备、变频器、发射设备和电源系统等部分组成。



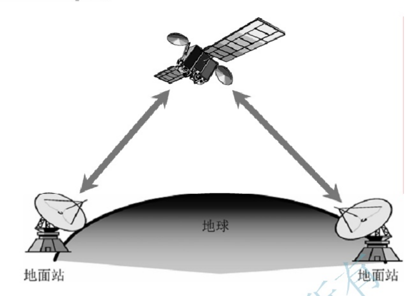


图 1.12　卫星通信系统组成示意图

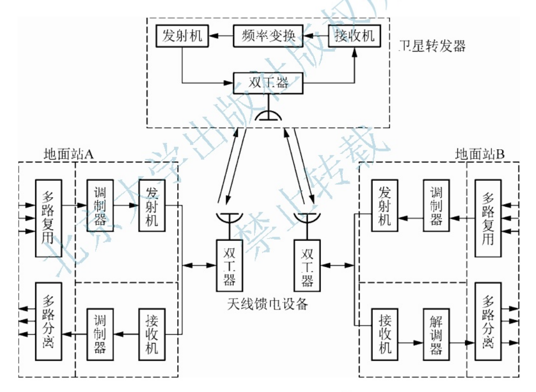


图 1.13　卫星通信系统线路组成框图

卫星通信作为现代无线通信的重要手段之一，其主要特点如下。

(1) 覆盖地域广、通信距离远、通信机动灵活、不受地理条件限制。

(2) 工作频段高。卫星通信的工作频率使用微波频段(0.3～30GHz)，主要原因是卫星处于外层空间，地面上发射的电磁波必须穿透电离层才能到达卫星，而微波频段对电离层的穿透能力最强。

(3) 频带宽、通信容量大、传输业务类型多。卫星通信采用微波频段，可供使用的频段很宽，星上能源和卫星转发器的功率保证充分，随着新技术和新体制的不断发展，卫星通信的容量也越来越大，传输业务的类型也日趋多样化。

(4) 通信质量好，可靠性高。卫星通信的电磁波主要在自由(宇宙)空间传播，而宇宙空间基本是真空状态，信道传输特性十分稳定，而且电磁波通常只经过卫星一次转发，受噪声和地面环境条件影响较小，通信质量稳定可靠，通信可靠性可达 99.8%以上。

(5) 线路使用费用与通信距离无关。地面微波中继或光缆通信系统，其建设投资和维护使用费都随距离的增加而增加，而卫星通信的地面站至空间转发器这一区间并不需要投资，因此线路使用费与通信距离无关。

目前，卫星通信系统被广泛地应用于军事、气象、资源探测、侦察、宇宙通信、科学实验、业务广播、全球定位等领域。21 世纪的卫星通信将向更高频段、更大容量方向发展，发挥全球覆盖的优势，建立卫星全球个人移动通信，从支持商业电信服务为主到面向最终个人消费者，地面移动终端由车载和便携向手持机发展，手持机将支持卫星和蜂窝双模式或多模式自动漫游。卫星间的通信将采用速度快、频带宽、保密性强的激光通信，宽带和窄带数据的卫星系统平行应用，低地球轨道卫星、中地球轨道卫星与静止轨道卫星及其他通信系统密切结合，优势互补，卫星网络与地面网络无缝衔接，地面终端设备将日益小型化，甚小天线卫星通信(VSAT)、卫星移动通信和卫星电视直播将会继续发展。



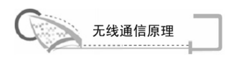
应用案例

动　中　通

动中通移动卫星通信系统(简称动中通)是指天线载体在移动中，天线伺服系统自动进行卫星跟踪，实现移动载体实时卫星通信的系统。动中通系统由卫星自动跟踪系统和卫星通信系统两部分组成，主要安装在移动物体上，如汽车、轮船等(图 1.14)，有自动寻星定星功能，通过动中通系统，车辆、轮船、飞机等移动的载体在运动过程中可实时跟踪卫星等平台，不间断地传递语音、数据、图像等多媒体信息，可满足各种军民用应急通信和移动条件下的多媒体通信的需要。例如，在地震、洪灾、雪灾等突发事件的实时报道时，有关受灾的最新情况大都采用动中通实现现场转播，从而能更快捷地为广大观众提供生动的电视画面。动中通车能在 20～100km/h 的行驶速度下通过卫星双向传送数字电视信号，保障运动载体在移动过程中不间断进行宽带多媒体卫星通信。动中通系统很好地解决了各种车辆、轮船等移动载体在运动中通过地球同步卫星，实时不断地传递语音、数据、高清晰的动态视频图像、传真等多媒体信息的难题，目前已广泛应用于部队、公安、武警、边防、消防、人防以及各级政府应急指挥通信平台的建立。



图 1.14　动中通



1.3.2 微波中继通信系统

微波中继通信是指利用微波作为载波并采用中继(接力)方式在地面上进行无线通信的过程。微波通信是在第二次世界大战后期，由美国贝尔研究所开始研究使用的一种无线电通信技术，微波通信经过几十年的发展，目前已经获得广泛的应用。微波通信用于地面上的长途通信时，需要采用中继(接力)接收并转发的传输方式才能完成信号从信源到信宿的传输任务。

微波通信的频率范围为 300MHz～300GHz，所对应的波长范围为 1m～1mm。具体的微波频段可细分为特高频(UHF)频段/分米波频段、超高频(SHF)频段/厘米波频段和极高频(EHF)频段/毫米波频段。由于卫星通信实际上也是在微波频段采用中继方式进行的通信，只是其中继站设在卫星上而已，因此，为了与卫星通信相区别，这里所说的微波中继通信仅是限定在地面上的微波通信。图 1.15 是远距离地面微波中继通信系统的示意图，其传输特点是在自由空间利用定向天线实现沿视距传输。由于受地形和天线高度的限制以及地球曲面和空间传播损耗的影响，空间两点间的传输距离一般为 30～50km，每隔 50km 左右就需要设置中继站，对信号进行接收、放大和转发，长距离微波通信可以经过几十次的中继而传输至数千 km 仍可保持高质量的通信。

微波中继通信主要用来传送长途电话信号、宽频带信号(如电视信号)、数据信号、移动通信系统基站与移动业务交换中心之间的信号等。

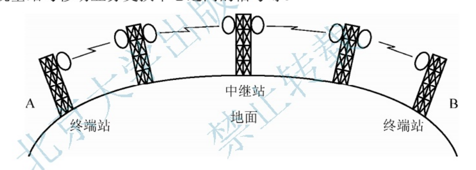


图 1.15 微波中继通信示意图

微波中继通信系统组成框图如图 1.16 所示，它主要包括微波终端站、微波中继站和天线馈线系统等部分。其中微波终端站包括微波收发信设备、调制解调设备、多路复用设备、电源设备和自动控制设备等；微波中继站主要完成信号的双向接收和转发，转发方式有微波转接方式、中频转接方式和基带转接方式；天线主要是完成馈线中传输的电磁能与自由空间传播的电磁波的相互转换，馈线则是电磁能量的传输通道。微波通信常用的天线形式有喇叭天线、抛物面天线、喇叭抛物面天线和潜望镜天线等。

微波中继通信的特点有以下几方面。

(1) 通信频带宽，传输容量大。占用的频带越宽，可容纳同时工作的无线电设备就越多，通信容量也就越大，一套短波通信设备一般只能容纳几条话路同时工作，而一套微波中继通信设备可以容纳几千甚至上万条话路同时工作，并可传输电视图像等宽频带信号。

(2) 受外界环境干扰影响小。微波通信具有良好的抗灾性能，对于水灾、风灾以及地震等自然灾害，微波通信一般都不受影响。

(3) 通信灵活性较大。微波中继通信采用中继方式，可以实现地面上的远距离通信，并且可以跨越沼泽、江河、湖泊和高山等特殊地理环境。在遭遇地震、洪水、战争等灾祸时，通信的建立、撤收及转移都比较容易，微波中断通信在这些方面比电缆通信具有更大的灵活性。

(4) 天线增益高、方向性强。由于微波具有直线传播特性，可利用微波天线将电磁波聚集成很窄的波束，使微波天线具有很强的方向性，以减少通信中的相互干扰。

(5) 投资少、建设快。微波中继通信线路可以节省大量的有色金属，建设费用低，建设时间也较短。

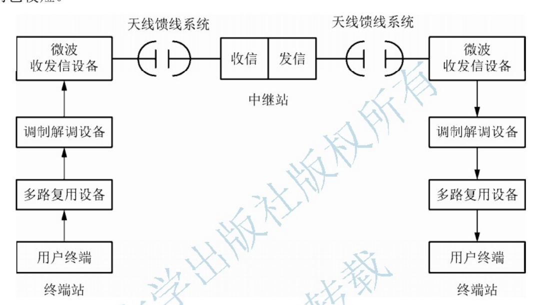


图 1.16　微波中继通信系统组成框图

1.3.3　短波无线通信系统

短波通信是无线电通信的鼻祖，多年来被广泛应用于航海、航空、军事、气象、外交、商业等部门，用以传输语言、数据、文字和图像等。短波通信是利用波长为 10～100m(频率为 3～30MHz)的电磁波进行的无线电通信，实际上，人们常把中波的高频频段 1.5～3MHz 也归到短波波段，所以现有的许多短波通信设备，其频率范围往往扩展到 1.5～30MHz。

现代短波无线通信系统一般由发信天线、集中发信台、遥控线路、集中收信台和收信天线等部分组成，短波无线通信系统组成示意图如图 1.17 所示。收发信台的主机一般由收发信道部分、频率合成器部分、功率放大器、逻辑控制部分、电源和一些选件部分组成，现代收发信机多了自适应选件，能借助收、发信道完成自动链路的建立。收发信天线完成输出的高频电流和电磁波之间的相互转换。

短波无线通信的特点如下。

(1) 无需中继站即可实现远距离的通信，因而建设和维护费用低，建设周期短。

(2) 通信设备简单，可以根据使用要求固定设置进行定点固定通信，也可以背负或装入车辆、舰船、飞行器中进行移动通信。



(3) 电路调度容易，临时组网方便、迅速，具有很大的灵活性。

(4) 短波通信对于自然灾害或战争的抗毁性较强。

(5) 通信设备体积小，容易隐蔽，便于改变工作频率以躲避干扰和窃听，破坏后容易恢复。

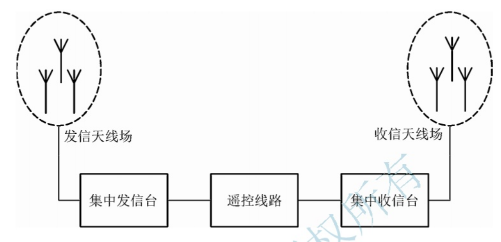


图 1.17 短波无线通信系统示意图

正是由于以上优点，短波通信一直被广泛用于军事、公安、边防、石油、矿山、渔业、森林防火、抗洪救灾、江河管理等部门，尤其适用于边远偏僻地区，山区、船岸之间的通信联络。在超短波通信距离有限、卫星通信受设备和话费制约的情况下，短波是很多用户解决远程移动通信的唯一手段。短波无线通信系统是国防通信网的重要组成部分之一，主要用于建立各级指挥员之间快速、直接指挥通信和陆、海、空三军之间的协同通信，并能保障方向不明、距离不等、位置不定的指挥员与作战部队之间的通信。

短波无线通信的主要发展方向如下。

(1) 采用实时选频技术，以选择最佳信道。由于短波电离层反射信道参数和各种干扰的时变性，工作频率的选择不能完全依赖于长期预报而要求实时预报，即利用实时选频技术找到电波不会穿出电离层、传播损耗小、多径时延小和无干扰或弱干扰的最佳工作频率。

(2) 采用各种自适应技术，以适应变参和干扰严重的短波信道。这些自适应技术包括数据速率的自适应、发射功率的自适应、对基带内窄带干扰的自适应(基带内频率捷变技术)、自适应均衡和使天线方向图的零点总对准干扰方向的自适应等技术。

(3) 采用新的分集接收技术、差错控制技术和压缩扩张技术等，进一步提高短波通信抗衰落和抗干扰的性能。

1.3.4 数字蜂窝移动通信系统

移动通信出现在半个世纪以前，20 世纪 80 年代以后得到了迅速发展，特别是随着蜂窝组网技术的完善和大容量系统的出现，移动通信已经成为发展速度最快、最受欢迎、最灵活方便的无线通信系统之一。数字蜂窝移动通信系统是目前全球应用最广的无线通信系统，因其无线覆盖小区的平面邻接拼图很像蜂窝而得名。数字蜂窝移动通信系统是将通信范围分为若干相距一定距离的小区，移动用户可以从一个小区运动到另一个小区，依靠终

端对基站的跟踪，从而使通信不中断。移动用户还可以从一个城市漫游到另一个城市，甚至到另一个国家与原注册地的用户终端通话。

数字蜂窝移动通信系统的基本组成如图 1.18 所示，该系统主要由基站(BS)、移动台(MS)和移动电话交换中心(MSC)几部分构成，通过移动电话交换中心进入公用有线电话网(PSTN)，从而实现移动电话与固定电话、移动电话与移动电话之间的通信。移动电话交换中心(MSC)是整个网络的核心，其主要功能是控制整个蜂窝系统的工作和对用户进行管理，还负责与本地电话网的连接、交换接续以及对移动台的计费等。MSC 配接有两个记录用户信息的数据库，分别为归属位置寄存器 HLR 和访问位置寄存器 VLR，它们主要用于配合 MSC 对移动用户进行管理。基站分布在每个小区，它包含控制单元、收发信机组、天馈系统、电源与数据终端等。基站主要负责本小区内移动用户与移动电话交换中心之间的连接，为移动用户提供接入网络的无线接口。每个基站都通过一个基站控制器(BSC)连接到 MSC，一个基站控制器可以控制多个基站。

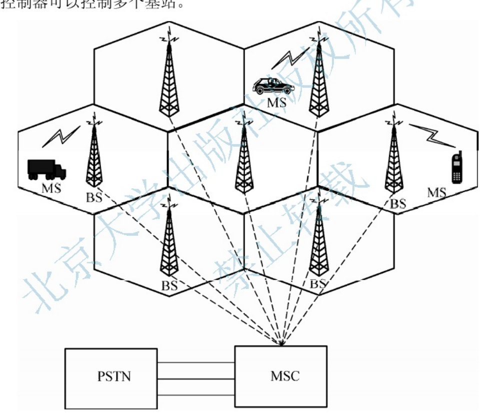


图 1.18　蜂窝移动通信系统的组成

数字蜂窝移动通信系统工作时，移动台通过无线接入信道与基站建立联系，如果移动台呼叫的是固定电话用户，则基站一方面将移动台的信号经 MSC 转接给 PSTN，另一方面也将来自 PSTN 的信号通过无线信道转接给移动台。当移动台从一个基站覆盖小区移动到另一个基站覆盖小区时，MSC 会控制基站将对移动台的服务从一个小区转移到另一个小区。每个基站的无线覆盖区都是一个正六边形的无线小区，无线小区邻接形成的几何图形形似蜂巢，所有无线小区相互邻接覆盖整个业务区域。为减小不同小区基站之间的同频干扰，需要使相邻小区的基站工作在不同的频率。蜂窝移动通信系统的设计建立在蜂窝概念

的基础上，蜂窝概念使频率复用成为可能，使得无线网络运营商可以无限次地重复使用无线电管理部门分配的有限频谱资源，从而可以设计出理论上用户容量无限大的蜂窝电话系统，并且可以根据需要不断扩充容量。

蜂窝移动通信具有用户容量大、频谱利用率高、通信质量好、业务种类多、易于保密、用户终端设备小巧轻便、成本较低以及便于与 ISDN、PSTN(公共电话交换网)、PDN(分组数据网)等网络互连、能向用户提供随时随地的全双工通信等优点。蜂窝电话已经从最初的以 TACS、AMPS 为代表的第一代模拟语音服务，经过目前广泛应用的以 GSM、CDMA 为代表的第二代数字蜂窝移动通信系统的数字语音和数据服务，过渡到今天的第三代(3G)移动通信系统的多媒体服务。目前，3G 已经商用化，而兼容各种移动通信新技术的第四代标准 4G 也正悄然来临，让我们不得不感叹“如今通信技术更新换代的速度快得像坐过山车啊！”我们将在后面的章节中对移动通信系统进行介绍。

1.3.5　无线局域网系统

无线局域网(Wireless Local Area Network，WLAN)是局域网的无线连接形式，它能使计算机在无线基站覆盖范围内的任何地点(包括户内、户外)发送、接收数据。最早的无线局域网络是 1971 年由夏威夷大学的研究人员开发的第一个基于封包式技术的无线电通信网络，被称为 ALOHNET 网络，这个无线网络包括了 7 台计算机，采用双向星型拓扑横跨 4 座夏威夷的岛屿，中心计算机放置在夏威夷大学主校区所在的 Oahu 岛上。该无线网络采用无线电台替代电缆线以克服由于地理环境因素而造成的布线困难问题，从而实现夏威夷大学各校区间共享主校区的大型计算机。如今的无线局域网可以让笔记本电脑、PDA(Personal Digital Assistant，个人数字助理)等移动终端设备摆脱布线的束缚，自由接入 Internet，从而可以减少基础设施的投入，解决了布线困难场所的接入问题。与有线网络相比，无线局域网具有安装便捷、使用灵活、可移动、易扩展等多方面的优点，可以在医院、商店、企业、学校等地区为集团用户提供服务，在机场、会议中心、展览中心、火车站、咖啡厅、酒店等地区为公众用户提供服务。目前，无线局域网已逐渐被广大用户作为一般的网络连接而广泛地应用于诸多行业，如金融证券、教育、大型企业、工矿港口、政府机关、酒店机场等。

无线局域网主要由无线网卡、无线接入点(AP)、无线网桥和客户终端等硬件设备组成(图 1.19)。无线网卡是以无线方式连接用户终端进行上网使用的计算机配件。无线网卡按照接口的不同可以分为台式机专用的 PCI 接口无线网卡、笔记本电脑专用的 PCMICA 接口网卡和 USB 无线网卡等(图 1.20)。无线接入点不仅包含单纯的无线接入点，也是无线路由器、无线网关等设备的统称，无线接入点能够将各个无线终端连接起来，主要是提供无线终端对有线局域网和有线局域网对无线终端的访问，在访问接入点覆盖范围内的无线工作站可以通过它进行相互通信。无线网桥是为使用无线进行远距离传输的点对点网间互联而设计的，它是一种在链路层实现局域网互联的存储转发设备。无线网桥可以用于连接两个或多个独立的网络段，这些独立的网络段通常位于不同的建筑内，相距几百米到几十千米，可用于固定数字设备与其他固定数字设备之间的远距离、高速无线组网。

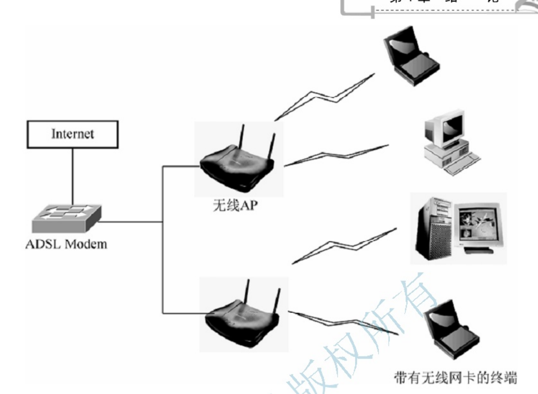


图 1.19 无线局域网的组成示意图



图 1.20 无线 AP 与无线网卡实例

无线局域网的拓扑结构有两种基本的类型：对等结构模式和基础结构模式。对等结构模式也称为自组织网络/移动自组网，是无线局域网的典型组网方式，这种拓扑的网络包含多个无线终端和一个服务器，均配有无线网卡，但不连接到接入点和有线网络，而是通过无线网卡进行相互通信。它主要用来在没有基础设施的地方快速而轻松地建立无线局域网。基础结构模式的拓扑结构是目前最常见的一种结构，这种结构包含一个接入点和多个无线终端，接入点通过电缆线与有线网络连接，通过无线电波与无线终端连接，从而可以实现无线终端之间以及无线终端与有线网络之间的通信，通过对这种模式进行复制，可以实现多个接入点相互连接的更大的无线网络。

IEEE 成立了 IEEE 802.11 工作组，主要负责制定无线局域网协议标准。1997 年 6 月，IEEE 推出了第一代无线局域网协议标准——IEEE 802.11 协议。该标准定义了物理层和介质访问控制子层(MAC)的协议规范，允许无线局域网及无线设备制造商在一定范围内建立可互操作的网络设备。1999 年 9 月，IEEE 提出 IEEE 802.11b 协议，用于对 IEEE 802.11 协议进行补充，之后又推出了 IEEE 802.11a、IEEE 802.11g 等一系列协议，从而进一步完

善了无线局域网规范。IEEE 802.11 系列协议标准在局域网领域内得到了广泛应用，是目前居于主导地位的无线局域网标准。

讲到无线局域网就不得不提 WiFi，目前 WiFi 似乎已被认为是无线局域网 WLAN 的代名词。WiFi 原为无线保真的缩写，其全称为 Wireless Fidelity，它是由 WiFi 联盟持有的一个无线网络通信技术的商标。WiFi 与蓝牙技术一样，同属于在办公室和家庭中使用的短距离无线通信技术，很多智能手机与笔记本电脑都支持 WiFi 上网。WiFi 目前可使用的标准有两个，分别是 IEEE 802.11a 和 IEEE 802.11b。如今，基于 WiFi 技术的无线局域网已经日趋普及，存在 WiFi 网络的公众场合，解决了运营商的互联互通、高收费、漫游性等问题。

无线局域网技术的研究和应用是目前无线通信领域的热点，家庭和小型办公网络用户对移动连接的需求，推动了无线局域网的迅速发展，其未来发展方向包括覆盖范围更广、支持速率更快、网络更智能、更易管理、解决漫游切换问题、应用多层次的安全策略提供不同等级的安全解决方案等方面。

1.3.6　无线传感器网络系统

无线传感器网络(Wireless Sensor Network，WSN)系统是当前在国际国内备受关注的、涉及多学科高度交叉、知识高度集成的前沿热点研究领域，它综合了传感器、嵌入式计算、现代网络及无线通信、分布式信息处理等技术。无线传感器网络是由部署在待监测区域内的大量的廉价的微型传感器节点组成，通过无线通信的方式形成的一个多跳的、自组织的无线网络系统。传感器、感知对象、观察者构成了无线传感器网络的 3 个基本要素。无线传感器网络的任务是协作地感知、采集和处理被该传感器节点组成的网络覆盖区域中的信息，经过一定的初级处理并发送给对待测信息感兴趣的观察者。人们可以通过无线传感器网络直接感知客观世界，从而极大地扩展了现有网络的功能和人类认识世界的能力。

无线传感器网络起源于美国，最初是 1978 年由国防部高级研究计划署(DARPA)在卡内基—梅隆大学发起的分布式传感器研讨会。具有代表性的项目包括：1993—1999 年间由美国国防高级研究计划署(DARPA)资助、加州大学洛杉矶分校(UCLA)承担的 WINS 项目，1999—2001 年间由 DAPRA 资助、UC Berkeley 承担的 Smart Dust 项目。传感器网络已经引起国内外科技界和工业界的高度重视，美国的《技术评论》将无线传感器网络列为第一项未来新兴技术，美国《商业周刊》和 MIT 技术评论在预测未来技术发展的报告中将无线传感器网络列为 21 世纪最有影响的 21 项技术和改变世界的十大技术之一。

无线传感器网络系统结构如图 1.21 所示，它包括传感器节点(Sensor 节点)、汇聚节点(Sink 节点)以及互联网或通信卫星和管理节点等。传感器节点通常是一个微型的嵌入式系统，具有感知、采集物理环境数据和处理控制数据的能力，并且还具有传统网络节点的终端功能和路由器功能。传感器节点除了进行本地信息采集和数据处理之外，还要对其他节点转发来的数据进行存储管理和融合。传感器节点实例如图 1.22 所示。汇聚节点通过无线方式连接传感器网络与 Internet 等外部网络，实现两种通信协议栈之间的通信协议转换，充当基站管理设备和传感器网络之间的通信员，可以发布基站管理节点的监测任务，并转发收集到的数据。管理节点一般就是普通的计算机系统，充当无线传感器网络服务器的角色，通过与传感器汇聚节点的信息传递来控制整个网络的数据和状态。无线传感器网络系

统的关键技术有网络协议、时间同步、定位技术、网络拓扑控制、数据融合、数据管理、无线通信技术、嵌入式操作系统、应用层技术等。

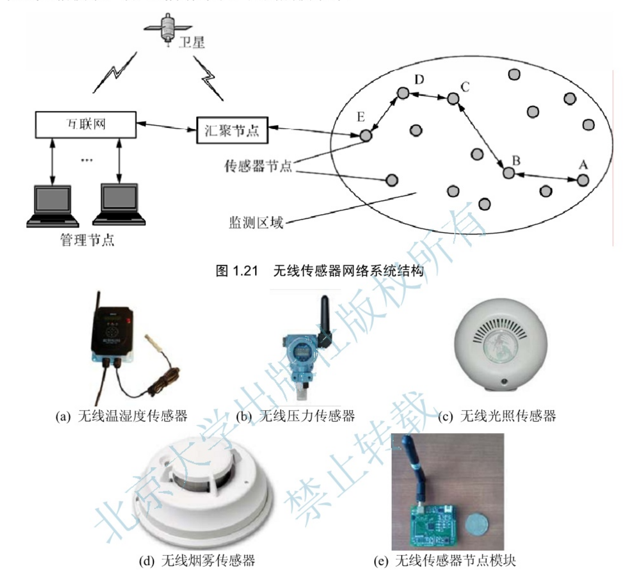


图 1.21 无线传感器网络系统结构

图 1.22 无线传感器实例

无线传感器网络作为信息技术中的又一个新的领域，它具有以下特点。

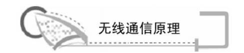
(1) 无线传感器网络中的传感器节点数量众多、分布密度较高，使得单位面积上所拥有的网络节点数量远远大于传统的无线网络，并且这些大量的传感器节点根本无须底层基础设施的支持就能够自动组织构成一定结构的网络。

(2) 无线传感器网络具备很强的容错能力。

(3) 无线传感器网络节点无中心，所有的节点都是地位平等，加入和离开节点都是随意的，没有严格的限制条件。

(4) 无线传感器网络节点与节点之间采用自组织的通信方式。

(5) 无线传感器网络中的传感器节点的电能是由节点所携带的电池直接供电的，因而节点携带能量十分有限。



(6) 无线传感器网络节点的体积和所携带能量有限，这决定了传感器节点只能具备有限的计算和存储能力以及相对距离较短的通信能力。

(7) 无线传感器网络的拓扑结构极易发生变化。

(8) 无线传感器网络以数据为中心，没有统一的通信协议平台。

无线传感器网络系统可以广泛地应用于军事战场、环境监测、空间探索、医疗健康、家庭生活、地震洪灾等临时紧急场合、工业生产以及智能农业等诸多方面，广阔的应用领域使得无线传感器网络成为人们生活中不可缺少的部分。无线传感器网络将物理世界和信息技术连接起来，具有改造社会的潜力。灵活自适应的网络协议体系、跨层设计、寻求系统节能策略、节点自动配置以及与其他网络的融合将是未来无线传感器网络发展的主要趋势。



应用案例

智能社区无线局域网

图 1.23 是智能社区系统实例，它是融合了无线传感器(WSN)的一个社区局域网。智能社区系统利用现代无线通信技术、传感技术、射频识别技术(RFID)、信息处理控制技术、多媒体技术和网络系统，将各个分离的设备(如基站、个人计算机、智能终端)、功能(如识别、数据传输)和信息(如环境检测测量)等集成到相互关联的、统一和协调的物联网系统之中，从而实现社区内各种信息的采集、处理、传输、显示和高度集成共享，实现社区集中、高效、便利的管理，实现社区和家庭各种电气设备和安防设备的自动化、智能化监控，实现社区生活与工作安全、舒适、高效。

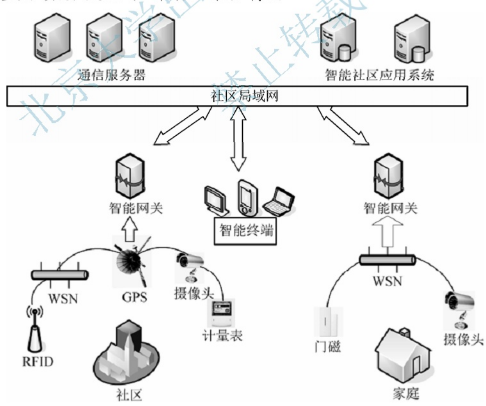


图 1.23　智能社区系统组成示意图

本 章 小 结

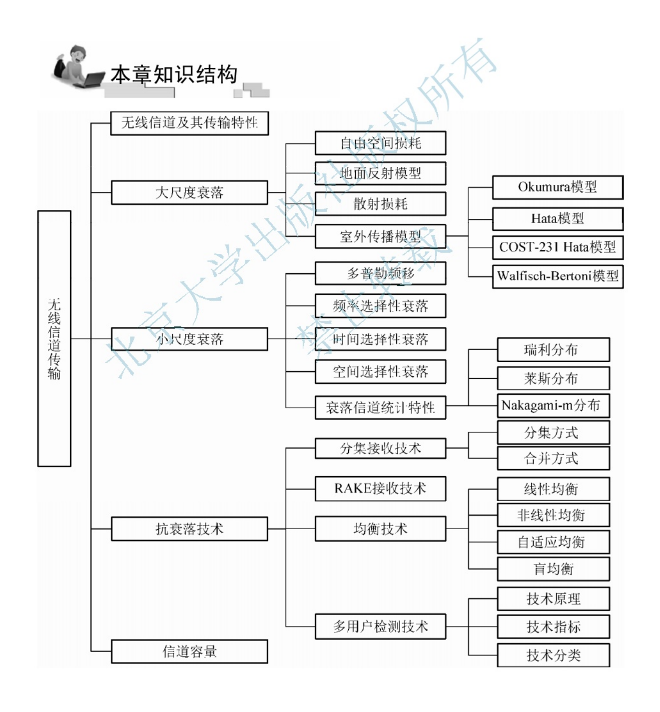
本章主要介绍了无线通信及其系统的基本组成、发展、特点，无线通信是指利用无线电磁波的辐射和传播，经过自由空间传送信息的通信方式，无线通信系统由信息源、发送设备、无线信道、噪声与干扰、接收设备、信宿六大基本部分组成。介绍了无线通信系统的分类和无线通信方式，根据不同的原则无线通信系统有不同的分类，根据不同的标准，通信方式同样也有不同的分类。本章还介绍了无线通信的主要技术指标，包括信道带宽、传输速率、频带利用率、差错率等。最后对几种典型的无线通信系统：卫星通信系统、微波中继通信系统、短波无线通信系统、数字蜂窝移动通信系统、无线局域网系统和无线传感器网络系统作了介绍，简单介绍了这几种典型无线通信系统的基本组成、发展和特点。

习 题

1.1 何谓无线通信？无线通信系统由哪几部分组成？无线通信的特点有哪些？
1.2 简述无线通信系统的分类和无线通信方式。
1.3 无线通信系统主要有哪些技术指标？
1.4 简述卫星通信系统的组成及其特点。
1.5 简述微波中继通信系统的组成及其特点。
1.6 简述短波无线通信系统的组成及其特点。
1.7 简述数字蜂窝移动通信系统的组成及其特点。
1.8 简述无线市话系统的组成及其特点。
1.9 简述无线局域网系统的组成、特点及其标准。
1.10 简述无线传感器网络系统的组成和特点。

第 2 章

无线信道传输



本章知识结构
无线信道传输
无线信道及其传输特性
大尺度衰落
自由空间损耗
地面反射模型
散射损耗
室外传播模型
Okumura模型
Hata模型
COST-231 Hata模型
Walfisch-Bertoni模型
小尺度衰落
多普勒频移
频率选择性衰落
时间选择性衰落
空间选择性衰落
衰落信道统计特性
瑞利分布
莱斯分布
Nakagami-m分布
抗衰落技术
分集接收技术
分集方式
合并方式
RAKE接收技术
均衡技术
线性均衡
非线性均衡
自适应均衡
盲均衡
多用户检测技术
技术原理
技术指标
技术分类
信道容量



本章教学目的与要求

- 了解无线信道及其特点
- 理解和掌握无线信道的传输特性
- 理解和掌握无线信道的大尺度衰落特性和小尺度衰落特性
- 理解和掌握几种室外传播损耗模型和自由空间传播损耗的计算
- 掌握频率选择性衰落、时间选择性衰落和空间选择性衰落的原理和主要参数
- 熟悉分集接收技术并掌握空间分集、频率分集、时间分集的原理
- 理解和掌握 RAKE 接收技术的原理
- 理解和掌握时域均衡技术的原理、结构和分类
- 理解和掌握多用户检测技术的原理、技术指标和分类以及各种检测技术的结构



引言

信道分为有线信道和无线信道，有线信道包括同轴电缆、对称电缆和光缆等传输介质，而无线信道有地波传播、短波电离层反射和微波视距中继等。有线信道的传输特性相对较稳定，如同轴电缆利用内部的金属介质传输信号，外部还有绝缘材料，信号在这样的信道中传输时各项参数基本是稳定的，而无线信道就不同了，无线信道大部分都是指自由空间，它的传输特性是随机变化的，并且受环境的影响非常大，信号在无线信道中传输会产生各种衰落现象，本章将主要介绍无线信道及其特点、无线信道对信号传输的衰落特性，包括大尺度衰落和小尺度衰落特性、传播路径损耗、衰落信道的类型以及有效的抗信道衰落技术。

【案例 2.1】无线信道的多径传播

无线信道是一种随时间、地理条件和其他外部因素而变化的传播环境，电磁波在发送端到接收端的无线信道中进行传输时必然要受到各种不同环境的影响。因此，在这样的环境中，电磁波的传播方式有很多种，图 2.1 给出了电磁波无线信道传输示意图，电磁波在传播过程中会受到建筑物、树木、植被以及起伏的地形的影响，引起电磁波的直射、反射、折射和散射等变化，因此接收机收到的无线信号可能来自不同的传播方向，经过不同的传播路径，这种现象称为多径传播。多径传播会使信号产生各种不同的衰落损耗，即电磁波传输的时候，也要付出相应的成本，这个成本就是传输损耗，无线传输损耗与很多因素有关，表达形式也多样，适用范围也各不相同。

【案例 2.2】基站

在移动通信系统中，基站是通过空中的无线信道与移动台相连的设备。大家常看到房顶上高高的天线，那就是基站的一部分。一个完整的基站包括无线发射/接收设备、天线和所有无线接口特有的信号处理部分。基站主要负责移动信号的接收、发送处理。一般情况下在某个区域内，多个子基站组成一个蜂窝状的网络，通过控制基站与基站之间的信号相互传送和接收来实现移动通信信号的传送，这个范围内的地区也就是我们常说的网络覆盖

面。如果没有了基站，就不可能完成手机信号的发送和接收。基站收发设备在基站控制器的控制下，完成基站的控制与无线信道之间的转换，实现手机通信信号的收发与移动平台之间通过空中无线传输及相关的控制功能。基站收发设备可对每个用户的无线信号进行解码和发送。基站使用的天线分为发射天线和接收天线，且有全向和定向之分，一般天线可有下列 3 种配置方式：发全向、收全向方式；发全向、收定向方式；发定向、收定向方式。图 2.2 分别给出了中国移动基站图、中国联通基站图和小灵通基站图。

图 2.1　电磁波无线信道传输示意图

(a) 中国移动基站　　(b) 中国联通基站　　(c) 小灵通基站

图 2.2　基站示意图

2.1　无线信道及其传输特性

2.1.1　无线信道

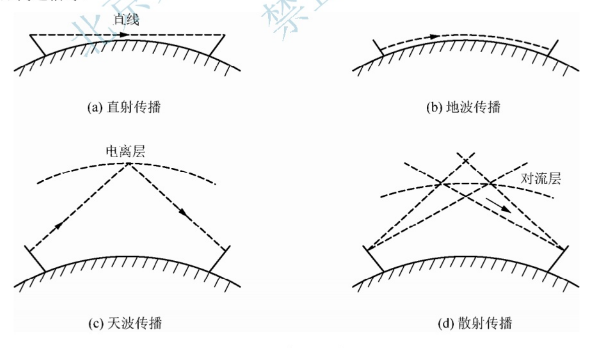
无线通信系统中位于发送设备和接收设备之间的无线传播路径是非常复杂的，电磁波从发射天线到接收天线可以有多种传播方式，如地球表面传播、空间波传播、对流层散射、电离层反射和地波绕射等。人们根据这些各具特点的传播方式，将电磁波归纳为 4 种主要类型(图 2.3)。

(1) 地面波。沿大地和空气的分界面传播的电磁波称为地面波或表面波，其传播途径主要决定于地面的特性，这种传播方式主要用于长波和中波远距离通信、军用短波和超短波小型电台的近距离通信等。

(2) 天波。天波也称电离层波，它是由天线向高空辐射的电磁波被大气上空的电离层反射而折回地面接收点的电磁波。地球大气层的高层存在着电离层，电磁波进入电离层时会受到折射而发生方向变化，由于电离层折射效应的积累，电磁波的入射方向会连续改变，最终会返回地面，因此电离层如同一面镜子会反射电磁波。天波传播主要用于短波远距离通信等。

(3) 空间波。空间波由直射波和反射波共同组成，直射波是由发射天线发送并直接到达接收点的电磁波，反射波是由地面或其他障碍物反射后到达接收点的电磁波。空间波传播主要用于超短波和微波视距范围通信等。

(4) 散射波。当大气层或电离层出现不均匀团块时，电磁波有可能被这些不均匀媒质向四面八方反射，使一部分能量到达接收点形成散射波。散射传播广泛用于超短波和微波远距离通信等。



(a) 直射传播　(b) 地波传播　(c) 天波传播　(d) 散射传播

图 2.3　电磁波的主要传播方式

无线信道的类型分为很多种，在通信工程中为了简化通信系统设计和性能分析，通常根据信道参数受外界影响时变化的快慢，将各种无线信道分为恒参信道和随参信道。恒参信道是指信道传输特性不随时间变化或变化很缓慢，如中长波地波传播信道、超短波及微波视距传播信道、人造卫星中继信道等。随参信道是指信道传输特性随时间随机快速变化，信号在传输中存在衰落、传输时延、多径效应等，如陆地移动信道、短波电离层反射信道、超短波流星余迹散射信道、超短波及微波对流层散射信道、超短波电离层散射以及超短波超视距绕射信道等都属于随参信道。

下面介绍几种较典型的随参信道。

1. 陆地移动信道

陆地移动通信工作频段主要在超高频(VHF)和甚高频(UHF)频段，电磁波传播特点是以直射波为主，由于城市建筑群和其他地形地物的影响，电磁波在传播过程中也会产生反射波、散射波以及它们的合成波。

在 VHF、UHF 移动信道中，电磁波传播方式主要有自由空间直射波、地面反射波、大气折射波、建筑物等，的散射波等，当移动台和基站天线在视距范围之内，这时电磁波传播的主要方式是直射波。电磁波除了直接传播外，遇到障碍物，例如，山丘、森林、地面或楼房等高大建筑物，还会产生反射或散射。电磁波在空间传播时，由于大气中介质密度随高度增加而减小，导致电磁波在空间传播时会产生折射、散射等。大气折射对电磁波传输的影响通常可用地球等效半径来表征。地球的实际半径和地球等效半径之间的关系为

$$k = \frac{r_e}{r_0} \tag{2-1}$$

式中，k 称为地球等效半径系数，$r_0 = 6370\text{km}$ 为地球实际半径，r_e 为地球等效半径。

例如，在标准大气折射情况下，地球等效半径系数 $k = 4/3$，此时地球等效半径为

$$r_e = kr_0 = \frac{4}{3} \times 6370 = 8493\text{km} \tag{2-2}$$

电磁波在传播过程中有一定的绕过障碍物的能力，称为绕射。当发射机和接收机之间不存在视距路径时，围绕阻挡体也产生波的弯曲，这也就是在障碍物后面有时仍能收到电磁波信号的原因。电磁波的绕射能力与电磁波的波长有关，波长越长，绕射能力越强；波长越短，则绕射能力越弱。

2. 短波电离层反射信道

电离层是距离地面 60～600km、处于电离状态的高空大气层，它主要是由于太阳辐射的紫外线、X 射线和宇宙射线等的作用使大气分子产生电离而形成的。当频率为 3～30MHz 的短波电磁波射入电离层时，由于折射现象会使电磁波发生反射并返回地面，从而形成短波电离层反射信道。

电离层存在 4 个密度不同的区域，即 D、E、F_1 和 F_2 这 4 层(图 2.4)。由于太阳辐射的变化，电离层厚度和高度会随昼夜及季节的不同而变化。在白天，由于太阳辐射强，所以 D、E、F_1 和 F_2 这 4 层都存在。在夜晚，由于太阳辐射减弱，D 层和 F_1 层几乎完全消失，因此只有 E 层和 F_2 层存在。电磁波射入电离层后，主要受到电离层的反射和吸收。反射



实现通信，而吸收则使电磁波能量受到衰减。电磁波的波长不同，电离层对它的反射和吸收程度也不同。一般而言，波长越长越容易产生反射，但电离层对其能量吸收也越大，反之，波长越短越不容易产生反射，电离层对其能量吸收也越小。由于 D、E 层电子密度小，不能形成反射条件，D、E 层对电磁波传输的影响主要是吸收电磁波，使电磁波能量损耗。F_2 层是反射层，其高度为 250～300km，一次反射的最大距离约为 4000km。

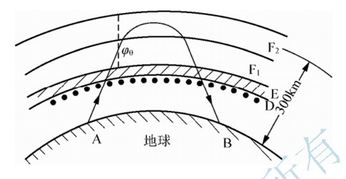


图 2.4　电离层结构示意图

短波电离层反射信道最主要的特征是多径传播，引起多径传播的主要原因有以下几种：①电磁波经电离层的一次反射和多次反射；②电离层反射层高度所形成的细多径；③电离层不均匀引起的漫射现象，主要影响是快衰落；④地球磁场引起的电磁波束分裂成寻常波与非寻常波。这 4 种情况如图 2.5 所示。

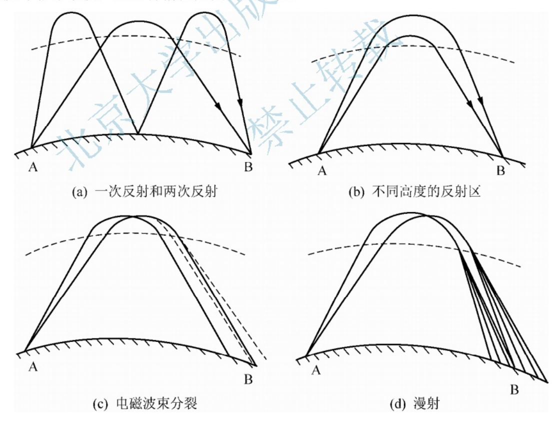
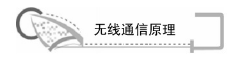


图 2.5　短波电离层反射信道的多径传输



3. 对流层散射信道

对流层散射是指电磁波在大气对流层的散射传播，对流层是大气的最低层，离地面高度 10～12km。在对流层中由于大气湍流运动等因素引起大气层的不均匀性，当电磁波射入对流层时，这种不均匀性就会引起电磁波的散射。对流层散射信道是一种超视距的传播信道，其一跳的传播距离约为 100～500km，可工作在超短波和微波波段。

2.1.2 无线信道的特点

(1) 传播的开放性。一切无线信道都是基于电磁波在空间传播来实现信息传播的，自由空间是一个开放的环境，所以电磁波在空间传播是开放式的信息传输，路径的空间约束性差，不确定性因素多，只能用随机模型来表示。

(2) 接收点地理环境的复杂性和多样性。如高楼林立的城市中心繁华区，一般性建筑物为主的近郊小城镇区，山丘、湖泊、平原为主的农村及远郊区，这些不同地形地貌建筑，以及气候和电磁干扰等情况都会对无线信道通信产生较大影响。

(3) 通信用户的随机移动性，如准静态的室内用户通信、慢速步行用户通信，高速车载用户通信。

(4) 无线信道的间歇性。由于室内人和物体的移动会对电磁波的传播造成时有时无的障碍，导致信道通断的间歇性，最终导致干扰正常通信。

2.1.3 无线信道的传输特性

无线通信系统的性能主要受到无线信道的制约，无线信道的传输特性直接影响到无线通信的质量。无线信道传输特性的研究主要针对以下 3 个问题：①某个特定频段和某种特定环境中，电磁波传播和接收信号的物理机制是什么；②从发射机到接收机，信号功率的路径损耗是多少；③接收信号的幅度、相位、多径分量到达的时间和功率是怎样分布的，统计特性如何。这样可以针对信号衰落特性，研究相应的抗衰落技术。在无线通信的不同传播环境下，本章主要讨论无线信道两方面的传输特性：大尺度衰落(大尺度路径损耗)和小尺度衰落。

大尺度衰落反映电磁波传播过程的路径损耗特性，决定无线通信系统的无线覆盖性能。路径损耗描述了发射机与接收机之间长距离(几百米到几千米)上的信号强度变化，所以也称为大尺度路径损耗。大尺度衰落决定了接收机与发射机相距一定距离时的平均接收信号电平，这个平均电平一般是随接收机所处位置不同而变化的。电磁波传播路径上遇到高大建筑物、树林、地形起伏等障碍物时会形成电磁场的阴影，产生阴影衰落或阴影效应。当接收机移动到这些阴影区域时，虽然接收机与发射机之间的距离没有变化，但平均接收信号电平会发生变化。另外，气象因素的变化也会影响信号传播的衰减特性。所以，在实际情况下，距离发射机相同距离处的实测接收信号的平均强度也是不相同的。大尺度衰落是无线网络规划设计中需要考虑的重要因素，在预测无线传播的大尺度路径损耗方面，人们已经完成了大量的研究工作，建立了适合不同传播环境的预测模型。

小尺度衰落是指接收机不移动，或者只在很短的距离上或很短的时间内移动时，引起



接收信号在短时间或短距离传播后其幅度、相位或多径时延快速起伏变化的特性，这种情况下对于大尺度路径损耗的影响可以忽略不计。由于在电磁波的传播路径上会存在各种不同的地形、地物，电磁波会受到各种不同地形、地物的阻挡而发生反射、散射等，因此接收机收到的无线信号可能来自不同的传播方向。小尺度衰落主要是由于无线信道的多径传播引起的，其他影响因素还有相对运动速度、周围物体的移动速度以及信号的传输带宽等。如陆地移动无线通信系统，因为发射机与接收机之间的相对运动造成传播环境的快速变化，信道具有时变特性，这些传播特性的变化率决定了信号的衰落速率。

小尺度衰落主要包括多径效应和多普勒效应等引起的信号包络的快速变化。由于接收者所处地理环境的复杂性，使得接收到的信号不仅有直射波的主径信号、还有从不同建筑物发射及绕射过来的多条不同路径的信号，而且它们到达时的信号强度、到达时间及到达时的载波相位都不一样，而所接收到的信号是上述各路径信号的矢量和，各路径之间也会产生相互干扰，这种干扰就称为多径效应。多径效应非常复杂，有时甚至收到的只是一些连续的反射波，而收不到主径直射波。多普勒效应是由于接收用户处于高速移动中，如车载通信时传播频率的扩散会引起多普勒效应，其扩散的程度与用户运动速度成正比。多普勒效应只产生在高速车载通信中，而对于通常慢速移动的步行和准静态的室内通信则不予考虑。

多径传播对无线通信系统性能和通信质量的影响巨大，当接收机天线处在深衰落位置点上时，甚至会造成通信中断。目前，在无线通信系统中已经采取多种抗衰落技术以减小多径传播的影响，如各种分集接收技术(时间分集、空间分集、频率分集)、均衡技术和匹配滤波技术等。本章在后面几节中将对大尺度衰落和小尺度衰落这两种衰落以及相应的抗衰落技术进行详细介绍。

2.2 大尺度衰落

大尺度衰落描述了长距离(几百米甚至更长)内接收信号的强度的缓慢变化。无线通信信道是一个完全开放式的信道，其传播损耗从宏观的大范围看，主要取决于传播的环境与条件。传播损耗不仅取决于传播距离，而且还与传播中的地形、地貌、传播的载波频率，以及收、发天线高度等密切相关。一般来说，大尺度衰落与发送天线和接收天线之间的距离成反比，且在不同的地区(如海边和内陆地区、城市和乡村)有不同的衰减因子。大尺度衰落可以由天线分集和功率控制得到补偿，因此掌握大尺度衰落对于无线通信传输技术和接收设备的选择有很大的意义。

2.2.1 自由空间传播损耗

自由空间传播是指从发射机到接收机之间没有任何影响传播路径的阻碍物体、反射物体和吸收物体的情形，介质是各向同性而且均匀的。电磁波的自由空间传播是发射机与接收机之间最简单的电磁波传播方式，卫星通信、空间通信和陆上视距通信都是自由空间传播。自由空间传播损耗主要指电磁波在理想的均匀的各向同性的介质中传播，由于传播路

径中没有阻挡，所以电磁波能量不会被障碍物吸收，也不会产生反射和折射现象，电磁波只存在能量扩散引起的传播损耗。

在自由空间传播方式下，由发射机发出的电磁波以球面波的形式在各向同性的介质(空气)中向四面八方传播。当发射机和接收机的距离较远时，到达接收机的电磁波可以近似为平面波。设发射天线的辐射功率为 P_t，将发射天线看作各向同性自由空间中的一个点，接收天线到发射天线的距离为 d，则接收天线的功率

$$P_r(d) = P_t G_t G_r \frac{\lambda^2}{(4\pi d)^2 L} \tag{2-3}$$

其中 G_t 为发射天线增益，其含义是电磁波传播方向上单位面积的电磁波功率相对于全向均匀传播时的单位面积的功率之比；G_r 为接收天线增益，其含义与 G_t 类似，表示电磁波到来的方向上单位面积的电磁波功率相对于全向均匀传播时的单位面积的功率之比；d 为接收天线与发射天线之间的直线距离，单位为 m，L 为与传播无关的系统损耗因子($L \geqslant 1$)，通常 L 取决于传输线衰减、滤波损耗和天线损耗，$L=1$ 则表明系统中无硬件损耗。设 A_r 是接收天线的有效面积，它与天线的物理尺寸和结构形式有关，也与波长有关，一般的 A_r 表示为

$$A_r = \frac{\lambda^2}{4\pi} \tag{2-4}$$

其中 λ 为电磁波的波长，单位为 m，与频率有关。

$$\lambda = \frac{c}{f} = \frac{2\pi c}{\omega_c} \tag{2-5}$$

其中 f 为工作频率，单位为 Hz；ω_c 为角频率，单位为 rad/s；c 为光速，单位 m/s。

自由空间传播路径损耗表示信号能量的衰减，通常以分贝(dB)的形式表示，定义为有效发射功率与接收功率之间的差值，单位为 dB 的正值，其定义式为

$$PL(\text{dB}) = 10\lg\frac{P_t}{P_r} = 10\lg\left[\frac{(4\pi d)^2 L}{G_t G_r \lambda^2}\right] \tag{2-6}$$

当发射天线增益和接收天线增益都等于 1 时，即 $G_t = G_r = 1$，且 $L=1$，则式(2-6)简化为

$$PL(\text{dB}) = 10\lg\frac{P_t}{P_r} = 10\lg\left[\frac{(4\pi d)^2}{\lambda^2}\right] \tag{2-7}$$

考虑到 $\lambda = c/f$，改用频率替代波长，并且 $c = 3\times10^8\,\text{m/s}$，则有

$$PL(\text{dB}) = 20\lg\frac{4\pi d}{\lambda} = 20\lg\frac{4\pi df}{c} = 32.44 + 20\lg d + 20\lg f \tag{2-8}$$

其中 d 为接收天线与发射天线之间的直线距离，单位为 km；工作频率 f 的单位为 MHz。由式(2-8)可以看出，自由空间传播损耗仅与距离 d 和工作频率 f 有关，距离越远损耗越大，工作频率越高损耗也越大。

【例 2.1】某发射机通过天线发射出去的功率为 10W，工作频率为 900MHz，接收天线增益为 3，发射天线的增益为 2，收发天线之间的距离为 10km，$L=1$。试求接收天线的接收功率和路径损耗。



解：已知 $P_t = 10\text{W}$，$G_t = 2$，$G_r = 3$，$d = 10\text{km}$，$L = 1$，$f = 900\text{MHz}$，

$$\lambda = \frac{c}{f} = \frac{3\times10^8}{900\times10^6} = \frac{1}{3}\text{m}$$

接收天线的接收功率为

$$P_r(d) = P_t G_t G_r \frac{\lambda^2}{(4\pi d)^2 L} = 10\times2\times3\times\frac{(1/3)^2}{(4\pi\times10\times10^3)^2} = 4.23\times10^{-10}\text{W}$$

自由空间传播路径损耗为

$$PL(\text{dB}) = 10\lg\frac{P_t}{P_r} = 10\lg\left(\frac{10}{4.23\times10^{-10}}\right) = 103.74\text{dB}$$

【例 2.2】设发射天线增益和接收天线增益均为 1，工作频率相同，传播距离 d_1、d_2 满足 $d_2 = 2d_1$，则这两种情况下传播损耗相差多少？

根据式(2-8)，传播距离 d_1、d_2 对应的传播损耗分别为

$$PL(\text{dB}) = 20\lg\frac{4\pi d}{\lambda} = 32.44 + 20\lg d_1 + 20\lg f$$

$$PL(\text{dB}) = 20\lg\frac{4\pi d}{\lambda} = 32.44 + 20\lg d_2 + 20\lg f$$

则两者的传播损耗差为

$$PL_2 - PL_1 = 20\lg d_2 - 20\lg d_1 = 20\lg(d_2/d_1) = 20\lg 2 = 6.02\text{dB}$$

可见，传播距离增加一倍，传播损耗会增加约 6dB。

同理，由例 2.2 可以看出，若传播距离相同而工作频率增加一倍，则传播损耗增加约 6dB。此外，若 $d_2 = 10d_1$，则两者的传输损耗差为

$$PL_2 - PL_1 = 20\lg(d_2/d_1) = 20\lg10 = 20\text{dB}$$

即传播距离增加到 10 倍时，传播路径损耗差为 20dB。图 2.6 给出了无线信道中自由空间传播损耗与频率和距离的关系。

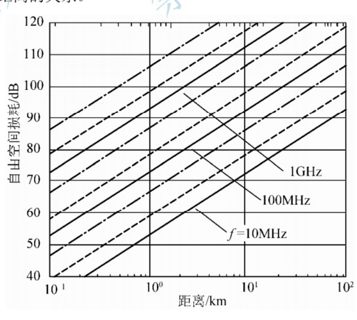


图 2.6　无线信道自由空间传播损耗与频率和距离的关系

2.2.2 地面反射(双线)模型

我们知道，在无线信道中，很少有单一路径传播的情况，一般会出现多径传播的现象。下面将讨论最简单的两条路径传播的模型，该模型仅考虑了空间的直射传播路径和地面的反射路径，分析模型如图 2.7 所示，该模型在预测几千米范围内的大尺度信号是非常准确的。

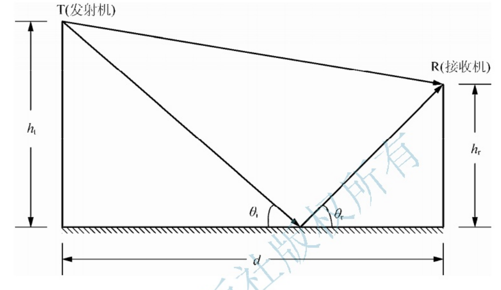


图 2.7 地面反射模型

假设发射机高度为 h_t，接收机高度为 h_r，发射机与接收机之间的水平距离为 d，可以得到与发射机水平距离为 d 处的接收机的接收功率为

$$P_r(d) = P_t G_t G_r \frac{h_t^2 h_r^2}{d^4} \tag{2-9}$$

这里只给出公式，具体推导分析过程可见相关参考书，路径损耗计算公式为

$$PL(\text{dB}) = 40\lg d - 10\lg G_t - 10\lg G_r - 20\lg h_t - 20\lg h_r \tag{2-10}$$

由式(2-9)和路径损耗公式(2-10)可见，当发射机与接收机之间距离很大时，接收机的接收功率随距离成 4 次方衰减，即 40dB/10 倍程距离，这比自由空间中 20dB/10 倍程的衰减要快得多，且此时接收功率与信号的频率无关。

【例 2.3】 GSM 系统的工作频率为 900MHz，基站天线高度为 50m，基站天线增益为 1，一个 GSM 移动台距离基站 5km，接收天线高度为 1.5m，接收天线增益为 1。试求①自由空间传播路径损耗；②使用双线地面反射模型的情况下的传播路径损耗。

解：已知 $P_t = 50\text{W}$，$G_t = G_r = 1$，$d = 5\text{km}$，$h_t = 5\text{km}$，$h_r = 1.5\text{m}$，$f = 900\text{MHz}$，则自由空间传播路径损耗为

$$PL(\text{dB}) = 32.44 + 20\lg d + 20\lg f = 32.44 + 20\lg 5 + 20\lg 900 = 105.52\text{dB}$$

若使用双线地面反射模型，则由双线地面反射模型路径损耗计算公式得传播路径损耗为

$$\begin{aligned}PL(\text{dB}) &= 40\lg d - 10\lg G_t - 10\lg G_r - 20\lg h_t - 20\lg h_r \\ &= 40\lg\left(5\times10^3\right) - 20\lg 50 - 20\lg 1.5 \\ &= 110.46\text{dB}\end{aligned}$$

【例 2.4】对于地面无线通信系统，使用双线模型分析时，平坦地区基站的覆盖范围是1km，如果使用卫星通信，则它的覆盖范围是多少？

解：在开阔地带，使用双线模型时传播路径衰减斜率为每 10 倍程距离增加 40dB。卫星通信的电磁波传播规律符合自由空间传播模型，传播路径衰减斜率为每 10 倍程距离增加 20dB。假设卫星通信的覆盖半径为 d，则根据题意应有

$$20\lg d = 40\lg 1000$$

即卫星通信的覆盖范围是 $d = 1000\text{km}$。

2.2.3　散射损耗

电磁波入射到尺寸远大于波长的光滑平面上时会产生反射，当反射平面粗糙时，电磁波会因散射而散布于所有的方向。因而，反射平面的粗糙程度会对反射效果产生很大的影响。另外，当电磁波遇到表面不是平面的物体时也会发生散射现象。在实际的移动通信系统中，树叶、街道标志和灯柱等都会引发散射。由于散射的能量散布于几乎所有的方向上，因此散射往往增加了接收信号的能量。电磁波发生散射的表面通常是粗糙不平的，设给定入射角 θ_i，则可得到表面平整度的参考高度 h_c

$$h_c = \frac{\lambda}{8\sin\theta_i} \tag{2-11}$$

其中 λ 为入射电磁波的波长。如果平面上最大的起伏高度 h 小于 h_c，则可以认为表面是光滑的；反之认为表面是粗糙的。通常，粗糙表面对反射效果的影响用散射损耗系数 ρ_s 来表示，Ament 提出在表面高度 h 的局部平均值服从高斯分布的情况下，ρ_s 为

$$\rho_s = \exp\left[-8\left(\frac{\pi\sigma_h \sin\theta_i}{\lambda}\right)^2\right] \tag{2-12}$$

其中 σ_h 是表面高度的标准方差。存在散射情况时，粗糙表面的反射系数 Γ 可以用散射损耗系数 ρ_s 来修正

$$\Gamma_{\text{rough}} = \rho_s \Gamma \tag{2-13}$$

2.2.4　室外传播路径损耗

在无线通信系统中，电磁波通常在不规则的地形环境中进行传播。在估计路径损耗时，应考虑传播路径上的地形地貌、建筑物、树木、植被等因素的影响，这些环境因素涉及了传播模型中很多变量，它们都起着重要的作用。从理论角度得出一个确切、完整的传播损耗公式难度很大。大多数传播模型是通过分析和试验相结合而获得的，通常在工程上大多采用一些经验公式与模型，这对于工程技术人员而言已基本上能满足工程上的估算要求。下面介绍几种典型的室外传播损耗模型(表 2-1)，这几种不同的室外传播损耗模型分别适用于不同的传播环境。

表 2-1　几种常见的室外无线传播模型

模型名称	适用范围
Longley-Rice 模型	40MHz～100GHz
Okumura 模型	150～1920MHz
Okumura-Hata 模型	150～1500MHz
COST-231 Hata 模型	扩展到 2GHz
Walfisch-Bertoni 模型	800～2000MHz

1. Longley-Rice 模型

Longley-Rice 模型是统计模型，它利用地貌地形的路径几何学和对流层的绕射性，预测了在自由空间中由于地形的非规则性造成的中值传输衰落。该模型适用于频率为 40MHz～100GHz、路径长度为 1～2000km、不同种类地形中点对点的通信系统，并且该模型已有了通用的计算机程序，以计算大尺度传输损耗。对于给定的传输路径，程序以传输频率、路径长度、天线高度、地面有效半径和地面导电常数等参数作为输入，可以预测电磁波通过不规则地形时的传输损耗。Longley-Rice 模型有两种模式：当地形路径数据很详细时，特定路径参数就很容易被确定，这种预测方式为点到点预测；如果地形地貌数据不够准确，可以利用 Longley-Rice 模型估计特定参数的值，这样的预测方式为区域预测。Longley-Rice 模型的缺点在于没有考虑接收机附近的环境因素，也没有考虑多径的影响。

2. Okumura 模型

Okumura 模型是 Okumura 根据日本大量测试数据统计出的以曲线图表示的传播模型。该模型广泛应用于城区信号预测，适用频率在 150～1920MHz(可扩展到 3000MHz)，适用距离 1～100km，要求基站天线的高度在 30～1000m。Okumura 开发了一套在准平滑城区、基站天线高度 h_{te} 为 200m、移动台高度 h_{re} 为 3m 的自由空间中值损耗 A_{mu} 曲线，并且移动台和基站均使用垂直全方向天线。Okumura 模型以准平坦地形大城市区的中值场强或路径损耗作为参考，对其他传播环境和地形条件等因素分别以校正因子的形式进行修正。该模型给出的是统计结果，在运用时必须考虑地形因子，它给出的测试结果显示郊区的损耗变化率较大，当接收天线较低时这一点更为明显。

应用 Okumura 模型计算路径损耗时，先要确定自由空间路径损耗，然后从所给曲线中读出中值损耗值，并加入代表地形的修正因子，Okumura 模型路径损耗的计算公式为

$$PL_{50}(\mathrm{dB}) = PL_{\mathrm{F}} + A_{\mathrm{mu}}(f,d) - G(h_{\mathrm{te}}) - G(h_{\mathrm{re}}) - G_{\mathrm{AREA}} \tag{2-14}$$

其中 PL_{50} 为传播路径损耗 50%处的值，即传播路径损耗中值，PL_{F} 是自由空间传播损耗，$G(h_{\mathrm{te}})$ 是基站天线高度增益因子，$G(h_{\mathrm{re}})$ 是移动台天线高度增益因子，G_{AREA} 是环境增益，A_{mu} 和 G_{AREA} 均是频率的函数，天线高度增益与天线的类型和形状无关。Okumura 模型已经给出了曲线，可以直接使用。图 2.8 和图 2.9 分别显示了 A_{mu} 和 G_{AREA} 的值。



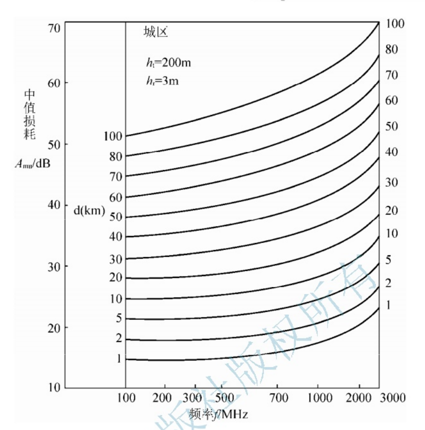


图 2.8　在准平滑地域上的自由空间中值损耗

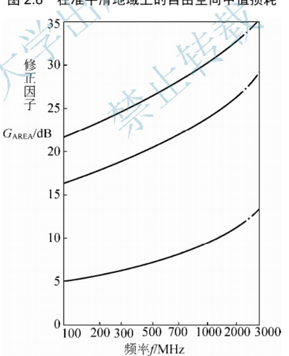
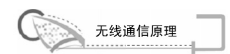


图 2.9　不同地形的修正因子 G_{AREA}

$G(h_{te})$ 和 $G(h_{re})$ 也有相应的公式

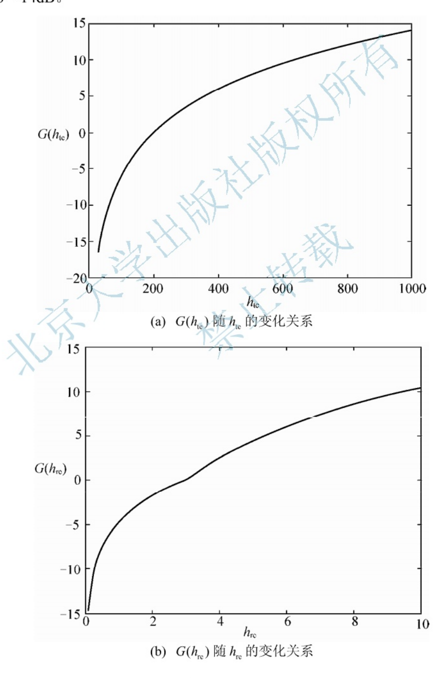
$$G(h_{te}) = 20\lg\left(\frac{h_{te}}{200}\right), \quad 30\text{m} < h_{te} < 1000\text{m} \tag{2-15}$$



$$G(h_{re})=\begin{cases}10\lg\left(\dfrac{h_{re}}{3}\right), & h_{re}\leqslant 3\text{m}\\ 20\lg\left(\dfrac{h_{re}}{3}\right), & 3\text{m}<h_{re}<10\text{m}\end{cases} \tag{2-16}$$

$G(h_{te})$ 和 $G(h_{re})$ 随 h_{te} 和 h_{re} 变化的关系分别如图 2.10(a)和图 2.10(b)所示。

Okumura 模型实用性较强，为成熟的蜂窝和陆地移动无线通信等系统的路径损耗预测提供了最精确的解决方法。但该模型完全建立在实验数据的基础上，没有理论的依据，对城区和郊区快速变化的反应很慢，不太适合地形变化太剧烈的地区，预测和实测的路径损耗偏差为 10～14dB。



(a) $G(h_{te})$ 随 h_{te} 的变化关系

(b) $G(h_{re})$ 随 h_{re} 的变化关系

图 2.10 天线高度增益因子

3. Okumura-Hata 模型

Okumura-Hata 模型也是一种广泛使用的传播模型，它是根据 Okumura 曲线图归纳所得到的经验公式。其频率范围为 150～1500MHz，主要用于 900MHz，适用于小区半径大于 1km 的宏蜂窝系统的路径损耗预测，基站高度超过其周围的建筑物。Okumura-Hata 模型以市区传播损耗为计算标准，其他情况需要在此基础上进行修正。Okumura-Hata 模型市区中值路径损耗经验公式为

$$PL_{50}(\text{市区})(\text{dB}) = 69.55 + 26.16\lg f_c - 13.82\lg h_{te} - a(h_{re}) + (44.9 - 6.55\lg h_{te})\lg d \quad (2\text{-}17)$$

其中 f_c 为载波频率；h_{te} 为发射天线有效高度，范围在 30～200m；h_{re} 为接收天线有效高度，范围在 1～10m；d 为发射机与接收机之间的距离，单位为 km；$a(h_{re})$ 为有效移动天线高度修正因子，是覆盖区大小的函数，其数值取决于环境。对于中小城市，移动天线高度修正因子为

$$a(h_{re})(\text{dB}) = (1.1\lg f_c - 0.7)h_{re} - (1.56\lg f_c - 0.8) \quad (2\text{-}18)$$

对于大城市有

$$a(h_{re})(\text{dB}) = 8.29(\lg 1.54 h_{re})^2 - 1.1, \quad f_c \leqslant 300\text{MHz} \quad (2\text{-}19)$$

$$a(h_{re})(\text{dB}) = 3.2(\lg 11.75 h_{re})^2 - 4.97, \quad f_c > 300\text{MHz} \quad (2\text{-}20)$$

在郊区的路径损耗，Okumura-Hata 模型经验公式修正为

$$PL_{50}(\text{dB}) = PL_{50}(\text{市区}) - 2\left[\lg(f_c/28)\right]^2 - 5.4 \quad (2\text{-}21)$$

在农村的路径损耗，Okumura-Hata 模型经验公式修正为

$$PL_{50}(\text{dB}) = PL_{50}(\text{市区}) - 4.78(\lg f_c)^2 + 18.33\lg f_c - 40.94 \quad (2\text{-}22)$$

当覆盖距离超过 1km 时，Okumura-Hata 模型和 Okumura 曲线所得的结果很接近。Okumura-Hata 模型适用于大区制移动系统，不适合覆盖距离不到 1km 的个人通信系统。

4. COST-231 Hata 模型

欧洲科学和技术研究协会(EURO-COST)组成 COST-231 工作委员会，提出了 Okumura-Hata 的扩展版本，即 COST-231Hata 模型，并将 Okumura-Hata 模型扩展到 2GHz。COST-231Hata 模型路径损耗的计算公式为

$$PL_{50}(\text{市区}) = 46.3 + 33.9\lg f_c - 13.82\lg h_{te} - a(h_{re}) + (44.9 - 6.55\lg h_{te})\lg d + C_M \quad (2\text{-}23)$$

其中 C_M 为大城市中心校正因子。在中等城市和郊区，$C_M = 0\text{dB}$；在市中心，$C_M = 3\text{dB}$。COST-231Hata 模型中参数的取值范围见表 2-2。

表 2-2　COST-231Hata 模型中参数的取值范围

参数名称	参数符号	取值范围
载波频率	f_c	1500～2000MHz
基站天线有效高度	h_{te}	30～200m
移动台天线有效高度	h_{re}	1～10m
收发天线距离	d	1～20km



COST-231Hata 模型和 Okumura-Hata 模型主要的区别在于频率衰减的系数不同，COST-231Hata 模型的频率衰减因子为 33.9，而 Hata 模型的频率衰减因子为 26.16。另外，COST-231Hata 模型还增加了一个大城市中心衰减因子 C_M，大城市中心地区路径损耗增加 3dB。

5. Walfisch-Bertoni 模型

由 Walfisch 和 Bertoni 合作开发的 Walfisch-Bertoni 模型考虑了自由空间损耗、沿路径传播的绕射损耗以及屋顶和建筑物高度的影响。Walfisch-Bertoni 模型的传播示意图如图 2.11 所示。

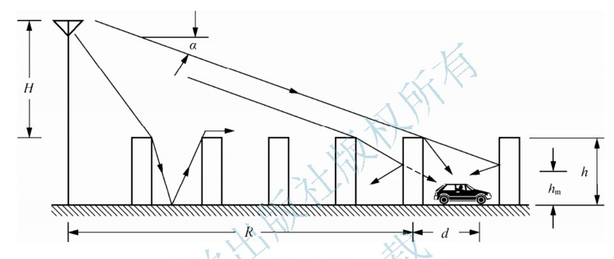


图 2.11 Walfisch-Bertoni 模型的传播示意图

该模型路径损耗 S 为 3 个因子的积

$$S = P_0 Q^2 P_1 \tag{2-24}$$

其中 P_0 为全向天线的自由空间路径损耗，因子 Q^2 为基于建筑物屋顶的信号衰减，P_1 是从建筑物屋顶到街道的基于绕射的信号衰减。Walfisch-Bertoni 模型的路径损耗用分贝的形式描述为

$$S(\text{dB}) = PL_0 + PL_{\text{rts}} + PL_{\text{ms}} \tag{2-25}$$

其中 PL_0 为自由空间损耗，PL_{rts} 为从建筑物屋顶到街道的绕射和散射损失，PL_{ms} 为建筑物的多屏绕射损耗。Walfisch-Bertoni 模型中参数的取值范围见表 2-3。

表 2-3 Walfisch-Bertoni 模型中参数的取值范围

参数名称	参数符号	取值范围
载波频率	f_c	800～2000MHz
基站天线有效高度	h_{te}	4～50m
移动台天线有效高度	h_{re}	1～3m
收发天线距离	d	0.2～5km

2.3 小尺度衰落

小尺度是描述短距离或短时间内接收信号强度的快速变化，无线信道的主要特征是多径传播，这些多径使得接收信号的幅度发生急剧变化，产生了小尺度衰落。无线信道的小尺度衰落主要有频率选择性衰落、时间选择性衰落和空间选择性衰落。这些衰落主要是由于无线信道的多径传播、移动台的运动和不同的散射环境，使得无线移动信道在时间上、频率上和角度上造成了色散而形成的。这些小尺度衰落特性分别与多径信号的时延扩展、多普勒扩展和角度扩展有关，这 3 种扩展又分别对应 3 组相关参数——相干带宽、相干时间和相干距离。

2.3.1 多普勒频移

当发射机与接收机之间存在相对运动时，接收机接收的信号频率与发射机发射的信号频率会不同，这种现象称为多普勒效应，接收频率与发射频率之差称为多普勒频移。当移动台以恒定速率 v 在长度为 d 、端点为 X 和 Y 的路径上运动时，收到发送点 S 发出的信号，如图 2.12 所示。电磁波从 S 点出发，在 X 点与 Y 点分别被移动台接收，所走的路径差为 $\Delta x_{\mathrm{i}} = d\cos\theta_{\mathrm{i}} = v\Delta t\cos\theta_{\mathrm{i}}$。这里 Δt 是移动台从 X 运动到 Y 所需的时间，θ_{i} 是 X 和 Y 处与入射角的夹角。由于发送点 S 距离很远，因此可假设 X 和 Y 处的 θ_{i} 相同，由路程差造成的接收信号相位变化值为

$$\Delta\varphi = \frac{2\pi\Delta x_i}{\lambda} = \frac{2\pi v\Delta t}{\lambda}\cos\theta_{\mathrm{i}} \tag{2-26}$$

其中 λ 为波长，由此可得出频率变化值，也就是多普勒频移 f_{d}

$$f_{\mathrm{d}} = \frac{1}{2\pi}\frac{\Delta\varphi}{\Delta t} = \frac{v}{\lambda}\cos\theta_{\mathrm{i}} \tag{2-27}$$

其中 v/λ 是 f_{d} 的最大值，它与入射角无关，因此最大多普勒频移为

$$f_{\mathrm{m}} = \frac{v}{\lambda} \tag{2-28}$$

由式(2-27)可知，多普勒频移与移动台运动速度及移动台运动方向与电磁波入射方向之间的夹角有关。若移动台朝向入射波方向运动，则多普勒频移为正(即接收频率上升)；若移动台背向入射波方向运动，则多普勒频移为负(即接收频率下降)。信号经不同方向传播，其多径分量造成接收机信号的多普勒扩散，从而使信号的带宽增加。

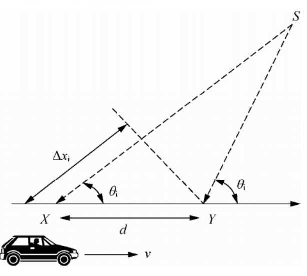


图 2.12 多普勒频移示意图

【例 2.5】无线通信基站发射机的载波频率为 1800MHz，分别计算移动台以 80km/h 的速度按下面 3 种情况运动时的接收信号载

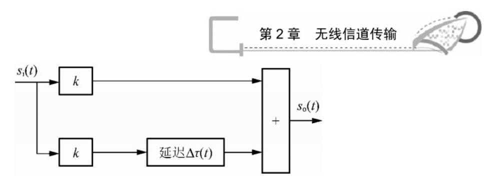


图 2.13　两条路径信道模型

由式(2-32)无线信道幅频特性可以看出，无线信道对信号的衰减程度与信号的频率有关，对于信号不同的频率成分，信道将有不同的衰减，这种特性就称为频率选择性衰落。对于固定的$\Delta\tau_i$，信道幅频特性如图 2.14(a)所示。显然，信号通过这种传输特性的信道时，信号的频谱将产生失真。特别是当信号的频谱宽度大于$1/\Delta\tau_i$时，这些频率分量甚至会被信道衰减到零，而造成严重的频率选择性衰落。另外，相对时延差$\Delta\tau(t)$通常是时变参量，这意味着传输特性中零点、极点在频率轴上的位置也随时间发生随机变化，这使传输特性变得更复杂，其特性如图 2.14(b)所示。

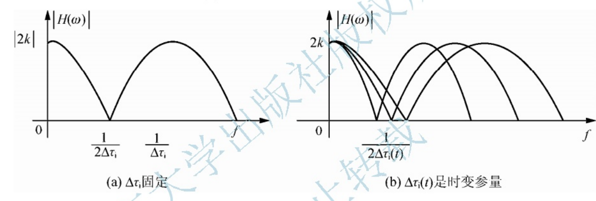


图 2.14　信道幅频特性

多径传播时的相对时延差通常用最大多径时延差来表征，设信道最大多径时延差为$\Delta\tau_m$，则定义多径传播信道的相干带宽为

$$B_c = \frac{1}{\Delta\tau_m} \tag{2-33}$$

它表示信道传输特性相邻两个零点之间的频率间隔。如果信号的频谱比相干带宽宽，则将产生严重的频率选择性衰落。为了减小频率选择性衰落，就应使信号的频谱小于相干带宽。在工程设计中，为了使接收信号不存在明显的频率选择性衰落，一般发送信号的带宽B应满足

$$B = \left(\frac{1}{5} \sim \frac{1}{3}\right) B_c \tag{2-34}$$

【例 2.6】设某无线信道的最大多径时延差为 2μs，为了避免发生频率选择性衰落，试估算在该信道上传输的数字信号的码元脉冲宽度。

解：该无线信道的相关带宽

$$B_c = \frac{1}{\Delta\tau_m} = \frac{1}{2\times10^{-6}} = 500\text{kHz}$$



根据工程设计要求，信号的带宽应选择为

$$B=\left(\frac{1}{5}\sim\frac{1}{3}\right)B_{\mathrm{c}}$$

所以数字信号的码元宽度

$$T_{\mathrm{s}}=\frac{1}{B}=(6\sim10)\mu\mathrm{s}$$

2.3.3 时间选择性衰落

由于移动台与基站之间的相对运动，接收信号在频域上产生多普勒频移现象，使得无线多径传播信道具有时变特性，则这种信道就具有了时间选择性衰落。时间选择性衰落会造成信号失真，这是由于发送信号还在传输的过程中，传输信道的特性已经发生了变化，即信号尾端时的信道特性与信号前端时的信道特性已经发生了变化。如果信号持续的时间比较短，在这个比较短的持续时间内，信道的特性还没有比较显著的变化，这时时间选择性衰落并不明显；当信号的持续时间进一步增加，信道的特性在信号的持续时间内发生了比较显著的变化时就会使信号产生失真。信号的失真随着信号的持续时间的增长而增加。发生时间选择性衰落时所对应的最小信号持续时间与最大多普勒频率的幅度呈反比的关系。

多普勒扩展和相干时间是用来描述无线信道的小尺度衰落时变特性的两个参数。多普勒扩展 B_{d} 是频谱展宽的一个测量值，它是移动无线信道时间变化率的一种度量。当发送频率为 f_{c} 的正弦信号时，接收信号的频谱即多普勒谱在 $f_{\mathrm{c}}-f_{\mathrm{m}}$ 至 $f_{\mathrm{c}}+f_{\mathrm{m}}$ 之间变化，其中 f_{m} 是最大多普勒频移。如果基带信号带宽 B 远大于多普勒扩展 B_{d}，则在接收端可忽略多普勒扩展的影响。

定义相干时间 T_{c} 是用来划分时间选择性和非时间选择性衰落信道的参数，而实际工程设计中对于相干时间有两种定义方法。其中一种定义相干时间 T_{c} 为多普勒扩展在时域上的表示，用于在时域上信道频率色散的时变特性，这种情况下相干时间 T_{c} 的定义式为最大多普勒频移 f_{m} 的倒数

$$T_{\mathrm{c}}\approx\frac{1}{f_{\mathrm{m}}} \tag{2-35}$$

相干时间的另一种定义为信道冲激响应维持不变的时间间隔的统计平均值，在这个时间间隔内，两个到达的信号的幅度衰落特性相同，如果两个到达信号的时间间隔大于此相干时间，则其幅度衰落特性的相关性很小，这种情况下相干时间 T_{c} 的定义式为信号包络相关度为 0.5 时的时间段长度，即相干时间的表达式近似为

$$T_{\mathrm{c}}\approx\frac{9}{16\pi f_{\mathrm{m}}} \tag{2-36}$$

显然，根据以上两种定义得到的相干时间 T_{c} 差别很大。为此，现代数字通信工程中普遍使用的相干时间的定义是取以上两种定义的平均，即

$$T_{\mathrm{c}}=\frac{0.423}{f_{\mathrm{m}}} \tag{2-37}$$

例如，移动台的移动速度为 30m/s，信道的载频为 2GHz，则相干时间为 1ms。

2.3.4　空间选择性衰落

由于无线通信中移动台和基站周围的散射环境不同，使得多天线系统中天线的点波束产生了角度色散，造成了同一时间、不同地点的信号衰落起伏不一样，即不同空间位置的天线经历的衰落特性不同的现象，这就是空间选择性衰落。角度扩展和相干距离就是描述空间选择性衰落的两个主要参数。

角度扩展 Δ (Azimuth Spread，AS)是由移动台或基站周围的本地散射体以及远端散射体引起的，它是用来描述空间选择性衰落的重要参数。所有的角度信息都与散射环境密切相关，移动环境中 3 种主要的散射体为移动台周围的本地散射体、基站周围的本地散射体和远端散射体。角度扩展描述了功率谱在空间上的色散程度，根据环境的不同，角度扩展在[0，360°]之间分布。角度扩展越大，表明散射环境越强，信号在空间的色散度越高；角度扩展越小，表明散射环境越弱，信号在空间的色散度越低，这也为智能天线的波束成形算法研究奠定了基础。

相干距离 D_c 是信道冲激响应保证一定相关度的空间间隔。相干距离是空间自相关函数的特有参数，它为衡量空间信号随空间相关矩阵变化提供了更普通的方法。在相干距离内，可以认为空间传输函数是平坦的，即信道可认为是非空间选择性信道。相干距离 D_c 与角度扩展和来波到达角有关，在天线的到达角相同的情况下，角度扩展越大，不同天线接收到的信号之间的相关性就越小，信号的空间选择性越严重；反之，角度扩展越小，天线之间的相关性就越大。同样，在角度扩展相同的情况下，信号的到达角越大，天线之间的相关性越大；信号的到达角越小，天线之间的相关性越小。因此，为了保证相邻两根天线经历的衰落不相关，在低散射环境下的天线间隔要比在高散射环境下的天线间隔长一些。

2.3.5　小尺度衰落信道类型

经过无线信道传播时，信号所经历的小尺度衰落类型取决于信号参数和信道参数之间的关系。根据无线信道的相干带宽、相干时间和发送信号的调制带宽、发送信号的码元持续时间的相对关系，可以将小尺度衰落信道分为平坦衰落、频率选择性衰落、快衰落、慢衰落、标量信道和矢量信道。

1. 平坦衰落信道

如果发送信号的带宽(调制带宽) B 远小于无线信道的相干带宽 B_c，或者发送信号的码元周期 T_s 远大于无线信道的时延扩展 σ_τ，即信道参数和发送信号参数满足如下关系

$$\begin{cases} B \ll B_c \\ T_s \gg \sigma_\tau \end{cases} \tag{2-38}$$

此时的无线传播信道就是平坦衰落信道，即非色散信道或幅度变化信道。通常，如果 $T_s \geqslant 10\sigma_\tau$，则可认为该无线传播信道是平坦衰落的。

显然，平坦衰落信道在发送信号的带宽内具有恒定的幅度增益和线性相位响应特性。因此，信号经平坦衰落信道传播后其频谱特性保持不变。在接收端，只有接收信号幅度随

信道的增益起伏而变化，接收端信号幅度的概率分布一般服从瑞利分布。信号经平坦衰落信道传播后的幅度变化会引起信号的深度衰落，衰落深度有时会超过 30dB。

2. 频率选择性衰落信道

当发送信号带宽大于无线信道的相干带宽，或者发送信号的码元持续时间小于无线信道的时延扩展，即信道参数和发送信号参数满足如下关系

$$\begin{cases} B > B_c \\ T_s < \sigma_\tau \end{cases} \tag{2-39}$$

则无线传播信道就表现为频率选择性衰落信道。通常，若 $T_s \leqslant 10\sigma_\tau$，该信道就可认为是频率选择性的。

信号经过频率选择性衰落信道传播后，由于信道的频率选择性，信号的不同频率成分将获得不同的增益，信号相位也发生变化，从而使发送信号产生频率选择性衰落，导致接收端信号发生失真。频率选择性衰落信道的建模比平坦衰落信道的建模更复杂，因为它可以看成是由多个具有不同时延的平坦衰落信道组合而成的。

3. 快衰落信道

当发送基带信号的码元周期大于无线信道的相干时间时，信道的冲击响应在码元持续时间内变化很快，这时的无线信道呈现出快衰落特性，称为快衰落信道。在快衰落信道中，接收端由于多普勒扩展引起的频率色散(也称时间选择性衰落)使接收信号失真，并且多普勒扩展越大，接收信号失真越严重。当发送信号的周期 T_s 大于信道的相干时间 T_c，且信号的带宽 B 小于相干带宽 B_d 时，信道冲激响应在码元周期内变化很快，从而导致信号失真而产生快衰，即快衰落信道满足下列条件

$$\begin{cases} T_s > T_c \\ B < B_d \end{cases} \tag{2-40}$$

从频域可看出，信号失真随发送信号带宽的多普勒扩展的增加而加剧。

4. 慢衰落信道

无线信道呈现快衰落特性还是慢衰落特性，取决于移动台、基站、空间反射体之间的相对运动的速度，以及基带信号发送的速率。当移动台、基站、空间反射体之间的相对运动速度较小时，信道的多普勒扩展就小。当信道的相干时间远远大于发送信号的周期，且信号的带宽 B 远远大于相干带宽 B_d 时，信道冲激响应的变化比要传送的信号码元周期低得多，信道呈现慢衰落特性，可认为该信道是慢衰落信道，慢衰落信道满足下列条件

$$\begin{cases} T_s \ll T_c \\ B \gg B_d \end{cases} \tag{2-41}$$

在慢衰落信道中，可认为信道参数在一个或多个信号码元周期内是稳定的。

需要注意的是，对于一个快衰落信道或慢衰落信道，它同时又可能是频率选择性衰落信道或平坦衰落信道。当慢衰落信道的相干带宽小于发送信号带宽时，慢衰落信道会同时呈现出频率选择性衰落，这时的信道称为频率选择性慢衰落信道；反之，当慢衰落信道的相



干带宽远大于发送信号带宽时，信道称为平坦慢衰落信道。同样，如果快衰落信道的相干带宽小于发送信号带宽，则此时的信道称为频率选择性快衰落信道，反之称为平坦快衰落信道。

5. 标量信道和矢量信道

在角度扩展不为零的情况下，随着多天线系统的引入，使得不同位置的天线经历了不同的衰落，从而产生角度色散，即空间选择性衰落；这样，信道信息从原来的时间、频率二维信息扩充到包含时间、频率、空间的三维信息。因此，根据信道是否考虑了空间选择性，把信道分为标量信道和矢量信道。包含时间、频率二维信息的信道称为标量信道；而包含时间、频率、空间三维信息的信道则称为矢量信道。因此，矢量信道的建模不仅需要考虑无线信道的衰落、时延等变量的统计特性，还必须考虑有关角度的统计特性，如到达角度(AOA)、离开角(AOD)和角度功率谱(PAS)等。

最后，将前面介绍的小尺度衰落信道进行总结，见表 2-4。

表 2-4　小尺度衰落信道分类

衰落信道类型	参　数	满足条件
平坦衰落信道	时延扩展 σ_τ	$B \ll B_c$ 且 $T_s \gg \sigma_\tau$
频率选择性衰落信道	相干带宽 B_c	$B > B_c$ 且 $T_s < \sigma_\tau$
快衰信道	多普勒扩展 B_d	$T_s > T_c$ 且 $B < B_d$
慢衰信道	相干时间 T_c	$T_s \ll T_c$ 且 $B \gg B_d$
标量信道	角度扩展 Δ	SISO(单天线)系统
矢量信道	相干距离 D_c	$\Delta \neq 0$ 的多天线系统

2.3.6　衰落信道的统计特性

1. 瑞利分布

瑞利分布常用来描述平坦衰落信号的接收包络或独立多径分量接收包络的时变统计特性。由于移动无线信道具有的多径传播特性，发射机发送的信号经过多条不同的路径到达接收端，假设多径信道一共有 N 条路径，且各条路径对信号的衰落是相互独立的，则接收机接收到的信号是多个路径的合成波。设发送信号为

$$s(t) = A\cos\omega_c t \tag{2-42}$$

经多径传播后得到的合成波为

$$r(t) = \sum_{i=1}^{N} a_i(t)\cos\left[\omega_c\left(t - \tau_i(t)\right)\right] \tag{2-43}$$

式中，$a_i(t)$ 为第 i 条路径到达接收端的信号振幅，$\tau_i(t)$ 为第 i 条路径的传输时延，传输时延可以转换为相位的形式，即

$$r(t) = \sum_{i=1}^{N} a_i(t)\cos\left[\omega_c t + \varphi_i(t)\right] \tag{2-44}$$

其中 $\varphi_i(t) = -\omega_c\tau_i(t)$，上式可进一步变换为

$$r(t) = \sum_{i=1}^{N} a_I(t)\cos\omega_c t - \sum_{i=1}^{N} a_Q(t)\sin\omega_c t \tag{2-45}$$



许多无线信道如散射信道、移动信道都存在多条传播路径，接收信号的幅度往往服从瑞利分布，因此这种快速衰落也常被称为瑞利衰落。

2. 莱斯分布

莱斯分布是由于在瑞利分布的基础上，存在一条直射路径的影响而造成的。当存在视距传播时，小尺度衰落的包络 $a(t)$ 服从莱斯分布，即

$$f(a)=\frac{a}{\sigma^2}\exp\left(-\frac{a^2+A^2}{2\sigma^2}\right)I_0\left(\frac{aA}{\sigma^2}\right),\quad 0\leqslant a<\infty \tag{2-55}$$

其中 A 为主信号幅度的峰值，$I_0(x)$ 为第一类修正贝赛尔函数。通常用莱斯因子 K 来描述莱斯分布，其定义为主信号的功率与多径分量方差之比，即 $K=A^2/2\sigma^2$ 或用 dB 表示为 $K(\mathrm{dB})=10\lg(A^2/2\sigma^2)$。当 $A\to0$，$K\to-\infty$，即主信号幅度减小时，莱斯分布转变为瑞利分布。因此，瑞利分布是莱斯分布的一个特例，莱斯分布是瑞利分布的一个扩展。

3. Nakagami-m 分布

Nakagami-m 分布是由 Nakagami 在 20 世纪 60 年代提出的，他通过基于现场测试的实验方法，用曲线拟合得到近似分布的经验公式。Nakagami-m 分布能描述大量的衰落环境，包括瑞利分布和莱斯分布，对于无线信道的描述有很好的适应性。若信号包络服从 Nakagami-m 分布，其概率密度函数为

$$f_A(a)=\frac{2m^m a^{2m-1}}{\Gamma(m)\Omega^m}\exp\left(-\frac{ma^2}{\Omega}\right),\quad a\geqslant0 \tag{2-56}$$

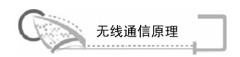
其中 $m=E(a^2)/var(a^2)$ 为不小于 $1/2$ 的实数，$\Omega=E(A^2)$，$\Gamma(m)=\int_0^\infty x^{m-1}\mathrm{e}^{-x}\mathrm{d}x$ 为伽马函数。显然，Nakagami-m 分布的概率密度函数与 m、Ω 两个参数有关，当 $m=1$ 时，Nakagami-m 分布退化为瑞利分布；当 m 较大时，Nakagami-m 分布接近于高斯分布。

2.4　抗衰落技术

无线信道是随机时变信道，如陆地移动信道、短波电离层反射信道等无线信道引起的衰落特性会严重影响接收信号的质量，降低通信系统的性能。为了提高无线信道中信号传输的质量，必须采用有效的抗衰落措施，通常采用的抗衰落通信技术包括分集接收技术、RAKE 接收技术、扩频技术、信道编码技术、抗衰落性能好的调制解调技术、功率控制技术、与交织结合的差错控制技术、均衡技术等。本节将介绍几种主要的抗信道衰落技术。

2.4.1　分集接收技术

分集接收技术是一种典型的抗多径衰落技术，并已在短波通信、移动通信系统等领域中得到广泛应用。分集接收的基本思想是指接收端按照某种方式使它收到的携带同一信息的多个信号衰落特性相互独立，并对多个信号进行特定的处理，有意识地分离多径信号并恰当合并以提高接收信号的信噪比来实现抗衰落。例如，人用两只眼睛和两只耳朵分别来



接收图像信号和声音信号就是典型的分集接收，一只眼睛肯定不如两只眼睛看得更清楚、更全面，一只耳朵的接收效果肯定不如两只耳朵的接收效果好。

分集接收技术包含有两重含义：一是分散接收，使接收机能够获得多个统计独立的、携带同一信息的衰落信号；二是集中处理，也就是接收机将收到的多个统计独立的衰落信号进行适当的合并，以降低衰落的影响，改善系统性能。

1. 分集方式

为了在接收端得到多个互相独立或基本独立的接收信号，一般可利用不同路径、不同频率、不同角度、不同极化、不同时间等接收手段来获取。因此，分集方式分为空间分集、时间分集、频率分集、极化分集、角度分集等多种方式。

1) 空间分集

空间分集又称天线分集，它是用多个天线接收同一个信号，其基本结构如图 2.16 所示，发送端采用一副天线，接收端采用多副接收天线。空间分集的接收机至少需要两副间隔一定距离的天线，为了使接收到的各信号相互独立，作为各接收天线之间的距离应满足

$$d > 3\lambda \tag{2-57}$$

其中 d 为各接收天线之间的距离，λ 为工作波长。

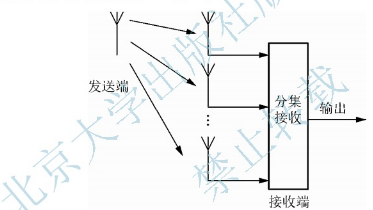


图 2.16 空间分集示意图

2) 频率分集

频率分集是利用多个载波信号传输同一信息，只要载波频率之间的间隔大到一定程度，则接收端所接收到信号的衰落是相互独立的。因此，载波频率的间隔应满足

$$\Delta f > B_c = \frac{1}{\Delta\tau_m} \tag{2-58}$$

其中 Δf 为载波频率间隔，B_c 为相干带宽，$\Delta\tau_m$ 为最大多径时延差。也就是说，当载波频率间隔大于相干带宽时，则可认为接收到信号的衰落是相互独立的。与空间分集相比，频率分集的优点是减少了天线的数目。其缺点是以增加带宽为代价要占用更多的频谱资源，有可能在发送端需多部发射机，使设备变得复杂。例如，在移动通信中，当工作频率在 900MHz 频段，典型的最大多径时延差为 $5\,\mu s$，此时载频间隔应满足

$$\Delta f \geqslant B_c = \frac{1}{\Delta\tau_m} = \frac{1}{5\times10^{-6}} = 200\text{kHz}$$

3) 时间分集

时间分集是将同一信号在不同的时间区间多次重发，只要各次发送的时间间隔足够大，则各次发送信号具有相互独立的衰落特性。在无线移动通信中，多普勒频移的扩散区间与移动台的运动速度及工作频率有关。因此，为了保证重复发送的数字信号具有独立的衰落特性，重复发送的时间间隔应满足

$$\Delta t \geqslant \frac{1}{2f_{\mathrm{m}}} = \frac{1}{2(v/\lambda)} \tag{2-59}$$

其中 f_{m} 为衰落频率，v 为移动台运动速度，λ 为工作波长。若移动台是静止的，则移动速度 $v=0$，此时要求重复发送的时间间隔 Δt 为无穷大。这表明时间分集对于静止状态的移动台是无效的。

时间分集的优点是减少了接收天线的数目；缺点是引入了信号处理时延，要占用更多的时隙资源，从而降低了传输效率。为了利用时间分集，必须在若干个相干时间周期内进行交织和编码。

4) 极化分集

极化分集是在发送端和接收端分别装有垂直极化和水平极化两副天线，这样就可以得到两路衰落特性不相关的信号。这一技术在蜂窝移动用户激增时，在改进链路的传输效率和提高容量方面有很明显的效果。

5) 角度分集

由于无线信道多径传播的影响，使得到达接收端的信号不可能是同方向的，角度分集是在接收端安装多个方向性尖锐的天线来得到不相关信号的一种分集方法。例如，在微波面天线上设置若干个照射器，产生相关性很小的几个波束。由于角度分集需要多个接收天线，并且接收天线分别指向不同的方向，因此角度分集是空间分集的一种变化形式。

此外，根据分集的位置，分集方式可以分为发射分集、接收分集和收发联合分集。根据信号传输的形式，分集方式可以分为显分集和隐分集。显分集是指构成明显分集信号的传输形式，多指利用多副天线接收信号的分集；隐分集是指分集作用隐含在传输信号的方式中，而在接收端利用信号处理技术实现信号的分集，一般只需一副天线。根据分集的目的，分集方式可以分为宏观分集(抗大尺度衰落)和微观分集(抗小尺度衰落)。宏观分集常用于蜂窝移动通信系统中，主要是克服由周围环境地形和地无差别而导致的阴影区引起的抗大尺度衰落；微观分集主要克服小尺度衰落。

2. 合并方式

合并是根据某种方式将接收端分集接收到的多个衰落特性相互独立的信号相加后合并输出，从而获得分集增益。从合并所处的位置来看，合并可以在检测器之前，即在中频和射频上进行合并，也可以在检测器之后，即在基带上进行合并，这两种形式分别如图 2.17 和图 2.18 所示。

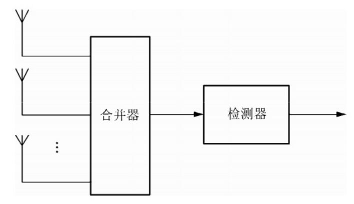


图 2.17　检测前合并技术

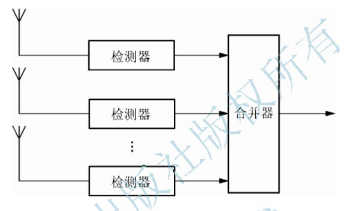


图 2.18　检测后合并技术

假设 N 个输入信号为 $r_1(t)$，$r_2(t)$，…，$r_N(t)$，则合并器的输出信号 $r(t)$ 为

$$r(t)=k_1r_1(t)+k_2r_2(t)+\cdots+k_Nr_N(t)=\sum_{i=1}^{N}k_ir_i(t) \tag{2-60}$$

其中 k_i 为第 i 路信号的加权系数。根据加权系数的不同，合并方式主要有选择式合并、等增益合并、最大比合并和切换合并。

1) 选择合并

选择合并是将 N 个分散接收的信号先送入选择逻辑，选择逻辑再从 N 个接收信号中选取信噪比最大的一个作为接收信号输出，其原理如图 2.19 所示。

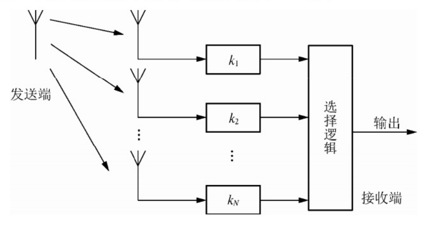


图 2.19　选择式合并原理图

选择式合并的平均输出信噪比为

$$\overline{r}_{\mathrm{M}} = \overline{r}_0 \sum_{k=1}^{N} \frac{1}{k} \tag{2-61}$$

合并增益为

$$G_{\mathrm{M}} = \frac{\overline{r}_{\mathrm{M}}}{\overline{r}_0} = \sum_{k=1}^{N} \frac{1}{k} \tag{2-62}$$

其中 $\overline{r}_{\mathrm{M}}$ 为合并器平均输出信噪比，$\overline{r}_0$ 为支路信号最大平均信噪比。对选择式合并，每增加一条分集路径，对合并增益的贡献仅为总分集支路数的倒数倍。

2) 等增益合并

等增益合并是将各支路信号以相同的支路增益进行直接相加，相加后的信号作为接收信号。等增益合并原理如图 2.20 所示。当加权系数 $k_1 = k_2 = \cdots = k_N$ 时，即为等增益合并。假设每条支路的平均噪声功率相等，则等增益合并的平均输出信噪比为

$$\overline{r}_{\mathrm{M}} = \overline{r}\left[1 + (N-1)\frac{\pi}{4}\right] \tag{2-63}$$

其中 $\overline{r}$ 为合并前每条支路的平均信噪比。合并增益为

$$G_{\mathrm{M}} = \frac{\overline{r}_{\mathrm{M}}}{\overline{r}} = 1 + (N-1)\frac{\pi}{4} \tag{2-64}$$

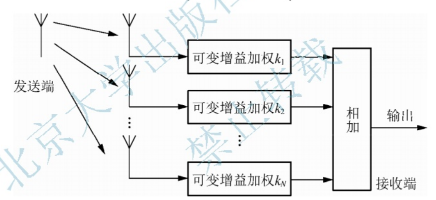


图 2.20　等增益合并原理

3) 最大比合并

最大比合并是最佳的分集合并方式，它可以获得最大的输出信噪比。最大比合并是各分集支路采用相同的衰落增益求权然后再合并，这个权值与本支路的信噪比成正比，信噪比越大，加权系数越大，对合并后信号贡献也就越大。若每条支路的平均噪声功率相等，则可以证明当各支路加权系数为 $a_k = A_K / \sigma^2$ 时，其中 A_K 为第 k 条支路信号的幅度，σ^2 为每条支路噪声平均功率，分集合并后的平均输出信噪比最大。最大比合并后的平均输出信噪比为

$$\overline{r}_{\mathrm{M}} = N\overline{r} \tag{2-65}$$

合并增益为

$$G_{\mathrm{M}} = \frac{\overline{r}_{\mathrm{M}}}{\overline{r}} = N \tag{2-66}$$

由此可见，合并增益与分集支路数 N 成正比。

比较上述 3 种合并方式，最大比合并性能最好，但其实现比较复杂；选择式合并实现最简单，但其性能也最差；等增益合并实现起来难度适中，其性能接近最大比值合并，是比较常用的一种合并方式。这 3 种合并方式的性能分析与比较如图 2.21 所示。

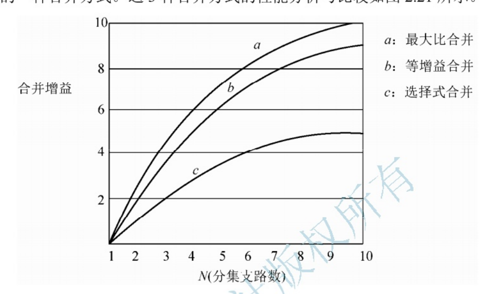


图 2.21　不同合并方式的增益比较

【例 2.7】某空间分集系统采用 4 重分集，分别计算选择式合并、等增益合并及最大比合并方法的合并增益。

解：由于分集支路数 $N=4$，则选择式合并增益为

$$G_{\mathrm{M}}=\frac{\overline{r}_{\mathrm{M}}}{\overline{r}_0}=\sum_{k=1}^{N}\frac{1}{k}=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}=\frac{25}{12}=2.08$$

等增益合并时的增益为

$$G_{\mathrm{M}}=\frac{\overline{r}_{\mathrm{M}}}{\overline{r}_0}=1+(N-1)\frac{\pi}{4}=3.36$$

最大比合并增益为

$$G_{\mathrm{M}}=\frac{\overline{r}_{\mathrm{M}}}{\overline{r}}=4$$

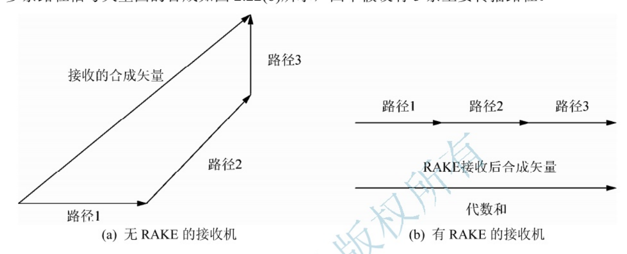
通过该例题可见，最大比合并增益最大，性能最好；选择式合并增益最小，性能最差。

2.4.2　RAKE 接收技术

RAKE 接收不同于传统的空间、频率与时间分集技术，它是一种典型的利用信号统计与信号处理技术将分集的作用隐含在被传输的信号之中的接收技术，RAKE 接收又被称为隐分集或带内分集。在 CDMA 扩频通信系统中，信道带宽远远大于信道的平坦衰落带宽，发送信号占用较宽的频谱资源，由于无线信道多径信号中含有可以利用的信源数据信息，所以 CDMA 接收机可以通过尽可能多地接收来自不同路径的信号，并按一定方式合并多径信号来改善接收信号的信噪比，从而克服多径衰落信道所造成的不良影响。其实 RAKE

接收技术就是通过多个相关检测器接收多径信号中的各路信号，并将各路信号分离开，相位校准，幅度加权，然后再将矢量和变为代数和。

RAKE 接收技术实际上是一种多径分集接收技术，若无 RAKE 接收机，则在接收端的多径传播信号的合成如图 2.22(a)所示，若采用扩频信号设计与 RAKE 接收的信号处理后，多条路径信号矢量图的合成如图 2.22(b)所示，图中假设有 3 条主要传播路径。



(a) 无 RAKE 的接收机　　(b) 有 RAKE 的接收机

图 2.22　多径信号矢量合成图

RAKE 接收机实现的原理框图如图 2.23 所示，它是专为 CDMA 系统设计的经典的分集接收器，其理论基础是当传播时延超过一个码片长度时，多径信号实际上可被看成是互不相关的。也就是说，在室外环境中，多径信号间的延迟通常较大，如果码片速率选择得当，那么 CDMA 扩频码的良好的自相关特性可以确保多径信号相互间表现出较好的非相关性。RAKE 接收机利用多个相关器分别检测多径信号中最强的 m 个支路信号，并按一定的规则进行合并，将矢量合并变为代数求和，以提供优于单路相关器的信号检测，并在此基础上进行解调和判决。

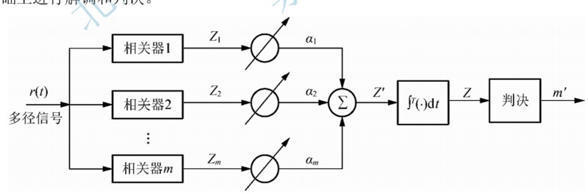


图 2.23　RAKE 接收机实现原理框图

设 CDMA 接收机有 m 个相关检测器，这些检测器的输出经过线性叠加并加权后，用来进行信号判决。假设相关器 1 与信号中的最强支路 m_1 同步，而另一相关器 2 与另一支路 m_2 同步，且 m_2 比 m_1 落后 τ_1。这里相关器 2 与支路 m_2 的相关性很强，而与 m_1 的相关性很弱。如果接收机中只有一个相关器，那么当其输出被衰落扰乱时，接收机无法作出纠正，从而会使判决器产生大量误判。而在 RAKE 接收机中，如果一个相关器的输出被扰乱了，

则还可以用其他支路来补救，并且通过改变被扰乱支路的权重，还可以消除此路信号的负面影响。由于 RAKE 接收机提供了对 m 路信号的良好的统计判决，因而它是一种克服衰落、改进 CDMA 接收的分集形式。在 CDMA 移动通信系统中，通常将 RAKE 接收与空间分集相结合，大大改善了传输条件，提高了系统性能。

在图 2.23 中，m 个相关器的输出分别为 Z_1，Z_2，…，Z_m，其权重分别为α_1，α_2，…，α_m，权重的大小是由各支路的输出功率或信噪比决定的。如果支路的输出功率或信噪比小，那么相应的权重就小。正如最大比合并分集方案一样，总的输出信号 Z' 为

$$Z' = \sum_{i=1}^{m} \alpha_i Z_i \tag{2-67}$$

其中权重 $\alpha_i (i = 1, 2, \cdots, m)$ 可用相关器的输出信号总功率归一化，其总和为 1，即

$$\alpha_k = \frac{Z_k^2}{\sum_{i=1}^{m} Z_i^2} \tag{2-68}$$

在研究分集合并时，曾有多种权重系数的生成方法。由于存在多址干扰，在这种情况下，多径信号强的支路在相关处理后未必能输出一个强信号，所以不能根据多径信号的强度确定某一支路的权重系数。只有基于相关器的实际输出来选择权重系数才能达到较好的 RAKE 接收性能。

2.4.3　均衡技术

无线通信系统中，由于无线信道的多径传播、衰落等影响会使接收端产生严重的码间干扰，码间干扰会使被传输的信号产生畸变，从而导致接收时产生误码。为了克服码间干扰，提高通信系统的性能，在接收端需采用均衡技术来有效地解决码间干扰问题。均衡是指对信道特性的均衡，也就是接收端滤波器产生与信道相反的特性，用来减小或消除因信道的时变多径传播特性引起的码间干扰。在无线通信系统中，通过接收端插入一种可调(或不可调)滤波器来校正或补偿系统特性，减小码间串扰的影响，这种起补偿作用的滤波器称为均衡器。图 2.24 所示为无线信道均衡示意图。



图 2.24　信道均衡示意图

实现均衡的途径有很多，目前主要是通过频域均衡和时域均衡两种途径来实现。频域均衡主要是从频域角度出发，使总的传输函数满足无失真传输条件，它是通过分别校正系统的幅频特性和群迟延特性来实现的。时域均衡主要是从时域响应考虑，使包含均衡器在内的整个系统总的冲激响应满足无码间干扰的条件。时域均衡实现起来比频域均衡更方便，性能一般也要优于频域均衡，故在时变的无线通信特别是移动信道中，几乎都采用时域均衡的实现方式。下面我们主要讨论时域均衡。

时域均衡器位于接收滤波器和抽样判决器之间(图 2.24)，它的基本设计思想是将接收



滤波器输出端抽样时刻上存在码间串扰的响应波形变换成抽样时刻上无码间串扰的响应波形。时域均衡在原理上分为线性均衡器和非线性均衡器两种类型，每一种类型均可分为多种结构，而每一种结构的实现又可根据特定的性能和准则采用多种自适应调整滤波器参数的算法。根据时域均衡器的使用类型、结构和算法的不同，对均衡器进行的分类如图 2.25 所示。下面将主要介绍几种典型的时域均衡器。

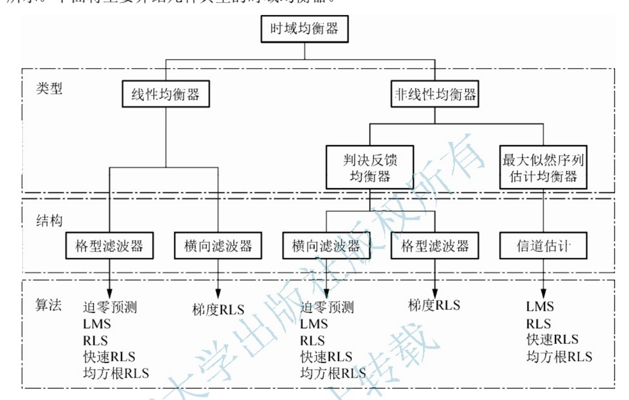


图 2.25　时域均衡器的分类

1.　线性均衡器

线性均衡器是指滤波器中的抽头加权系数为线性关系，抽头加权系数为非线性关系的则称非线性均衡器。在线性均衡器中，最常用的均衡器结构是线性横向(LTE)均衡器，如图 2.26 所示，它由 $2N+1$ 个抽头延迟线组成，以符号 C_{-N}，$C_{-(N-1)}$，…，C_0，…，C_{N-1}，C_N 表示各节的抽头增益系数，级与级之间的延时时间间隔均等于一个码元间隔 T_s，它将所收到的信号的当前值和过去值按滤波器系数进行线性叠加，并将生成的和作为输出。

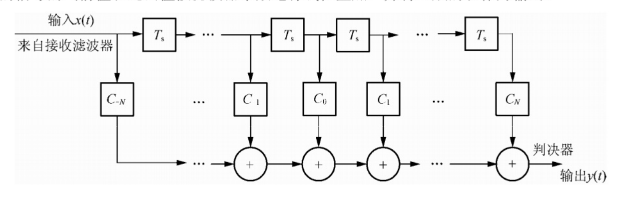


图 2.26　线性横向均衡器

根据线性系统的分析原理，线性横向均衡器的冲击响应为

$$h(t)=\sum_{n=-N}^{N}C_n\delta(t-nT_s) \tag{2-69}$$

当输入为 $x(t)$ 时，其输出为

$$y(t)=x(t)*h(t)=\sum_{n=-N}^{N}C_n x(t-nT_s) \tag{2-70}$$

因此，输出的第 k 个样值为

$$y_k=\sum_{n=-N}^{N}C_n x_{k-n} \tag{2-71}$$

式(2-71)说明，均衡器输出在第 k 个抽样时刻上的样值 y_k 将由 $2N+1$ 个 C_n 与 x_{k-n} 乘积之和来确定。显然，要想消除码间干扰，希望抽样时刻的输出不受前后码元的影响，即除 $k=0$ 外的所有 y_k 都要等于零。据此来求系数 C_n。经推导可以得到，当 $n\to\infty$，各抽头的增益系数为

$$C_n=\frac{T_s}{2\pi}\int_{-\pi/T_s}^{\pi/T_s}\frac{T_s}{X_{eq}(\omega)}\exp\left(jn\omega T_s\right)d\omega \tag{2-72}$$

时，才可以使码间干扰为零。式中 $X_{eq}(\omega)=\sum_n x_n\exp(jn\omega T_s)$，这个结果说明在理论上需要无穷多个滤波器才能实现，而这实际上是无法做到的，也是没有必要的，因此，可以采用某种准则作为依据从而将码间干扰减小到不妨碍判决时刻的正确性即可，常用的准则有峰值畸变准则和均方误差最小准则。

峰值畸变准则定义为

$$D=\frac{1}{y_0}\sum_{\substack{k=-\infty\\k\neq0}}^{\infty}|y_k| \tag{2-73}$$

其中将 $\sum_{\substack{k=-\infty\\k\neq0}}^{+\infty}|y_k|$ 表示除 $k=0$ 以外的各样值的绝对值之和，反映了码间干扰的最大值，y_0 是有用信号样值，所以峰值畸变 D 就是码间干扰最大值与有用信号样值之比。显然，设计时应使 D 最小，即对于完全消除码干串扰的均衡器而言，应有 $D=0$，对于码间干扰不为零的场合，希望 D 有最小值。

均方误差最小准则定义为

$$e^2=\frac{1}{{y_0}^2}\sum_{\substack{k=-\infty\\k\neq0}}^{\infty}{y_k}^2 \tag{2-74}$$

其物理意义与峰值畸变准则相似。

实际上不同的均衡器有不同的算法，对于环境不变的情况，均衡器的抽头一次调好就可以不再改变。而对于无线通信系统，由于环境的改变，均衡器的抽头往往需要能自动调整，这就是自适应均衡器。

2. 非线性均衡器

当信道中存在深度衰落而使信号产生严重失真时，线性均衡器会对出现深度衰落的频谱部分及周边的频谱产生很大的增益，从而增加了这段频谱的噪声，以致线性均衡器不能取得满意的效果，这时采用非线性均衡器处理效果比较好。常用的非线性算法有判决反馈均衡(DFE)、最大似然符号检测均衡及最大似然序列估计均衡(MLSE)。

1) 判决反馈均衡器

判决反馈均衡器(DFE)的结构如图 2.27 所示，它由两个横向滤波器和一个判决器构成，两个横向滤波器由一个前向滤波器和一个反馈滤波器组成，其中前向滤波器是一个一般的线性均衡器，前向滤波器的输入是接收序列，反馈滤波器的输入是已判决的序列。判决反馈均衡器根据接收序列预测前向滤波器输出中的噪声和残留的码间干扰，然后从中减去反馈滤波器输出，从而消除这些干扰，其中码间干扰是由硬判决之后的信号计算出来的，这样就从反馈信号中消除了加性噪声。与线性均衡器相比，判决反馈均衡器的错误概率要小。

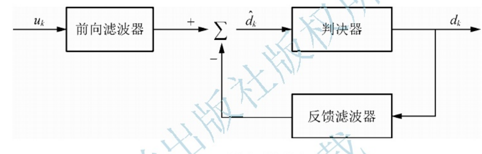


图 2.27　判决反馈均衡器

2) 最大似然序列估计均衡器

最大似然序列估计均衡器(MLSE)最早是由 Forney 提出的，它设计了一个基本的最大似然序列估计结构，并采用 Viterbi 算法实现。最大似然序列估计均衡器的结构如图 2.28 所示，最大似然序列估计均衡器通过在算法中使用冲击响应模拟器，并利用信道冲激响应估计器的结果，检测所有可能的数据序列，选择概率最大的数据序列作为输出。最大似然序列估计均衡器是在数据序列错误概率最小意义下的最佳均衡，这就需要知道信道特性，以便计算判决的度量值。

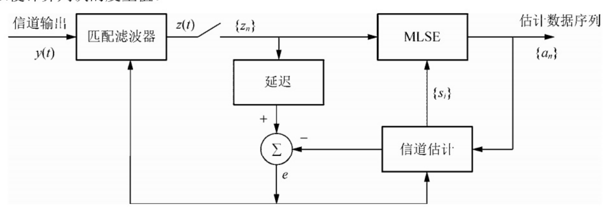
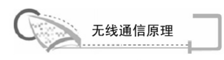


图 2.28　最大似然序列估计均衡器(MLSE)的结构



3. 自适应均衡器

由于无线信道的衰落具有随机性和时变性，这就要求均衡器必须能够实时地跟踪无线信道的随机时变特性，这种均衡器被称为自适应均衡器。均衡器的自适应均衡算法(Adaptive Algorithm)包括最小均方误差算法(Lowest Mean Square Error，LMSE)、递归最小二乘算法(Recursive Least Square，RLS)、快速递归最小二乘算法(Fast RLS)、平方根递归最小二乘算法(Square Root BLS)、梯度递归最小二乘算法(Gradient RLS)、最大似然比算法(Maximum Likelihood Ratio)、快速卡尔曼算法(Fast Kalman)等。在比较这些算法时，主要考虑算法的快速收敛特性、跟踪快速时变信道特性和尽可能小的运算量。

自适应均衡器有两个工作状态：训练方式自适应均衡工作状态和跟踪方式自适应均衡状态。在训练方式工作状态中，发送已知的测试序列启动均衡器并使之迅速收敛，完成抽头系数加权的初始化。在跟踪方式工作状态中，均衡器直接利用通信中传输的数字信号的判决形成误差信号，并依据自适应算法跟踪调节抽头系数，它能自动适应信道的随机变化，保持在最佳均衡准则下的信道均衡。在实际系统中，为了便于均衡器收敛，常在传输的数字信号中设置专门的训练序列。

4. 盲均衡器

近年来，盲均衡在无线通信和信号处理领域受到了普遍关注，盲均衡技术是指均衡器能够不借助训练序列，仅利用接收序列本身的先验信息来均衡信道特性。盲均衡利用发送信号的已知统计特性对信道进行自适应均衡，通过使均衡器输出的统计特性与接收信号的已知统计特性相匹配来调整均衡器的参数，使其输出序列尽量逼近发送序列。盲均衡的原理框图如图 2.29 所示，图中，$x(n)$ 为发送序列，$h(n)$ 为传输信道的冲激响应(包括发射滤波器、传输媒介和接收滤波器的综合作用)，$n(n)$ 为信道迭加噪声，$w(n)$ 为盲均衡器的冲激响应，$y(n)$ 为经过信道的接收序列，也是盲均衡器的输入序列，$\hat{x}(n)$ 为经过均衡后的恢复序列。

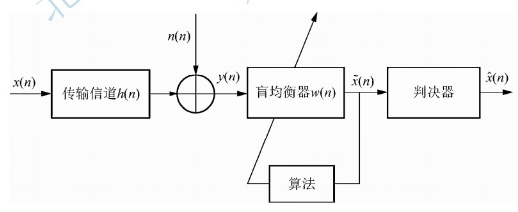


图 2.29 盲均衡的系统模型

目前，比较成熟的盲均衡算法有基于 Bussgang 技术的盲均衡算法、基于神经网络理论的盲均衡算法、恒模盲均衡算法、多模盲均衡算法、盲最大似然序列估计算法和基于高阶统计特性的盲均衡算法。盲均衡技术是一种新兴自适应均衡技术，它能有效地补偿信道的非理想特性，获得更好的均衡性，而且不必发送训练序列，有效地提高了信道效率。

2.4.4　多用户检测技术

1. 多用户检测技术的原理

自 1986 年美国学者 Verdu 提出最优多用户检测技术以来，多用户检测理论迅速成为通信理论界研究的热点，目前，多用户检测技术已成为第三代移动通信中宽带 CDMA 系统抗干扰的关键技术。多用户检测技术(MUD)基于干扰检测的思想，根据信息论中的最佳联合检测理论，利用检测所得到的信息来降低干扰对接收信号的影响，多用户检测技术能抗多址干扰和抗多径干扰。图 2.30 所示为多用户检测器的一般结构。

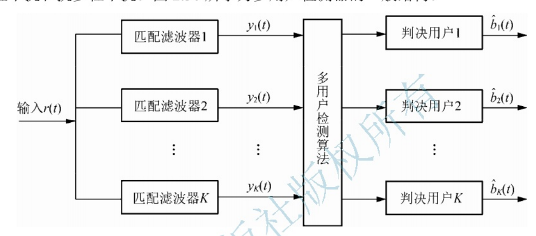


图 2.30　多用户检测器的一般结构图

在实际系统中，不论是多径衰落还是多址干扰，其本质上并不是纯粹无用的白噪声，而是具有强烈结构性的伪随机序列，而且各用户之间与各路径之间的相关函数已知，多用户检测器就是充分利用这些伪随机序列的已知信息和统计信息从而实现联合检测多个用户的信息，消除干扰，并从接收信号中估计出 $b_k(t)$。多用户检测技术是消除或减弱多址干扰和多径衰落干扰的有效手段，并能够消除 CDMA 中远近效应，简化 CDMA 系统中的功率控制，改善了系统性能、提高系统容量和增大小区覆盖范围。多用户检测技术的缺点是：系统的设备复杂度较高，处理延时较大，而且需要知道用户的一些信息，需要通过不断地进行信道估计来实现，估计的精度会直接影响检测器的性能。

2. 多用户检测器的技术指标

评价多用户检测器性能的技术指标主要有误码率、渐近有效性和抗远近能力。

1) 误码率

假设在加性高斯白噪声信道中不存在干扰用户时，具有能量 E_k 的单用户系统的误码率定义为

$$P_{k,\mathrm{su}}(\sigma)=Q\left(\sqrt{\frac{E_k}{\sigma^2}}\right) \tag{2-75}$$

其中 σ^2 为噪声方差。当存在多个用户时，由于多址干扰的存在，误码率会增大，此时用户 k 的误码率定义为

$$P_k(\sigma)=Q\left(\sqrt{\frac{e_k(\sigma)}{\sigma^2}}\right) \tag{2-76}$$

其中 $e_k(\sigma)$ 为用户 k 达到 $P_{k,\text{su}}(\sigma)$ 时所需的能量。

2) 渐近有效性

渐近多用户有效性简称为渐近有效性，它是衡量干扰用户对目标用户误码率影响的技术指标。多用户有效性定义为多用户系统到达单用户系统相同误码率所需能量与单用户系统所需能量之比，即

$$\eta_k(\sigma)=\frac{e_k(\sigma)}{E_k} \tag{2-77}$$

渐近多用户有效性定义为在高信噪比情况下多用户有效性的极限，即

$$\eta_k(\sigma)=\lim_{\sigma\to 0}\frac{e_k(\sigma)}{E_k}=\sup\left(\lim_{\sigma\to 0}\frac{P_k(\sigma)}{Q\left(\dfrac{\sqrt{rE_k}}{\sigma}\right)}=0\right)$$
$$=\frac{2}{E_k}\lim_{\sigma\to 0}\left(\sigma^2\log 2\frac{1}{P_k(\sigma)}\right) \tag{2-78}$$

渐近有效性 η_k 的取值介于 0 和 1 之间。

3) 抗远近能力

远近效应是指如果干扰用户比目标用户距离基站近很多，则干扰信号在基站的接收功率就会比目标用户信号的接收功率大很多，这样，传统接收机的输出中多址干扰分量就可能很严重，甚至会淹没目标用户的信号的现象。抗远近能力是用来描述多用户检测器抵抗远近效应能力的技术指标，它定义为在所有相关用户能量范围内测量到的最坏情况下的渐近有效性

$$\bar{\eta}_k=\inf_{\substack{E_j>0\\ j\neq k}}\eta_k \tag{2-79}$$

抗远近能力一般取决于扩频波形和解调器。

3. 多用户检测器的分类

多用户检测器的具体分类如图 2.31 所示。

下面简单介绍几种典型的多用户检测技术。

1) 最优多用户检测技术

最优多用户检测技术的原理是假设发送信号的先验概率相等，从接收信号 $r(t)$ 中找出发送序列，使其联合后验概率 $p(b\,|\,r(t):0<t<+\infty\}$ 最大。按照最大后验概率准则进行序列检验，有

$$b^*=\arg\{\max p(b\,|\,r(t))\} \tag{2-80}$$

最优多用户检测器是多用户检测器的最佳结构，可以达到最高的渐近有效性，理论上可以获得最小的错误概率。然而最优多用户检测器在工作时必须知道所有用户的信号幅度和相位，而且该方法具有较高的复杂度，其复杂度与用户个数成指数增长，因此当用户数过大时，该方法根本无法实现。最优多用户检测器包括同步最优多用户检测器和异步最优多用户检测器。



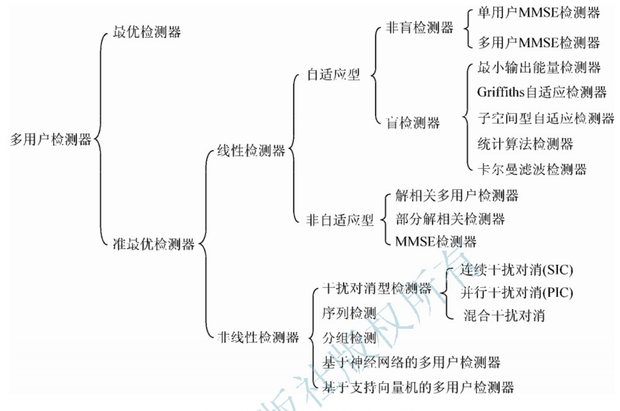


图 2.31　多用户检测器的具体分类

2) 线性多用户检测技术

线性多用户检测技术的基本原理是对匹配滤波器组的输出矩阵进行线性变换，然后再对产生的新的输出矢量进行判决，该类检测器的复杂度与用户数呈线性关系。线性多用户检测器主要包括解相关检测器、最小均方误差(MMSE)检测器、子空间斜投检测器、多项式展开检测器和自适应检测器等。

(1) 解相关检测器。解相关检测器的设计目的是为了去除所有用户之间的线性相关性，使不同用户的扩频波形相正交。图 2.32 所示为解相关检测器结构。

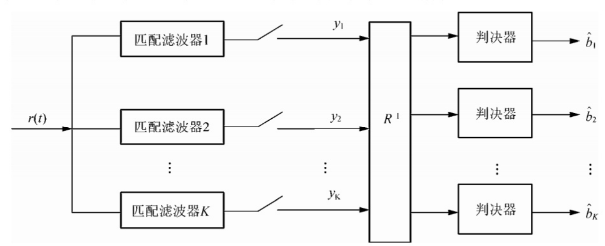


图 2.32　解相关检测器

解相关检测器接收信号模型

$$y = RAb + n \tag{2-81}$$

其中 $\boldsymbol{y} = (y_1, y_2, \cdots, y_K)^{\mathrm{T}}$，$\boldsymbol{b} = (b_1, b_2, \cdots, b_K)^{\mathrm{T}}$，$\boldsymbol{n} = (n_1, n_2, \cdots, n_K)^{\mathrm{T}}$，$\boldsymbol{A} = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_K \end{bmatrix}$，

发送信号矢量的检测可以看作求解线性方程组。令变换矩阵 $\boldsymbol{T}_{\text{dec}} = \boldsymbol{R}^{-1}$，则有

$$\hat{\boldsymbol{b}}_{\text{dec}} = \boldsymbol{R}^{-1}\boldsymbol{y} = \boldsymbol{Ab} + \boldsymbol{R}^{-1}\boldsymbol{n} = \boldsymbol{Ab} + \boldsymbol{z} \tag{2-82}$$

则每个用户的判决比特为

$$\hat{b}_k = \operatorname{sgn}\left[\left(\boldsymbol{R}^{-1}\boldsymbol{y}\right)_k\right] \tag{2-83}$$

线性变换后的噪声矢量的相关矩阵为

$$E(\boldsymbol{zz}^{\mathrm{T}}) = \boldsymbol{R}^{-1}E(\boldsymbol{nn}^{\mathrm{T}})\boldsymbol{R}^{-1} = \sigma^2\boldsymbol{R}^{-1} \tag{2-84}$$

第 k 个用户的误码率为

$$p_{\text{e,dec}}^{(k)} = p[\hat{b}_k = 1 \mid b_k = -1] = p[b_{\text{dec}} > 0 \mid b_k = -1] = p[z_{\text{dec}} > \mathrm{A}] = Q\left(\frac{A}{\sqrt{\sigma^2 R_{k,k}^{-1}}}\right) \tag{2-85}$$

其中 $R_{k,k}^{-1}$ 为矩阵 R^{-1} 的第(k,k)个元素。

解相关检测器完全消除了多址干扰，改善了系统性能，提高了容量，不需要估计信号幅度，且性能与干扰用户的容量无关，但其缺点是放大了噪声，也就是说完全抑制多址干扰是以提高噪声功率为代价的，且解相关检测器中的矩阵求逆很难实时运算。

(2) 最小均方误差(MMSE)检测器。最小均方误差(MMSE)检测器是考虑背景噪声和接收功率的线性检测器，以最小均方误差代价函数为目标，使发送比特矢量和匹配滤波器组输出矢量间的均方差最小化，即 $\arg\min\limits_{\boldsymbol{T}} E\left[\left|\boldsymbol{b} - \boldsymbol{Ty}\right|^2\right]$，将 $\boldsymbol{y} = \boldsymbol{RAb} + \boldsymbol{n}$ 带入上式，并求梯度可得线性变换矩阵为

$$\boldsymbol{T}_{\text{MMSE}} = \left(\boldsymbol{R} + \sigma^2\boldsymbol{A}^{-2}\right)^{-1} \tag{2-86}$$

进一步可得 MMSE 检测器的判决矢量

$$\hat{\boldsymbol{b}}_{\text{MMSE}} = \left(\boldsymbol{R} + \sigma^2\boldsymbol{A}^{-2}\right)^{-1}\boldsymbol{y} \tag{2-87}$$

每个用户的判决比特为

$$\hat{b}_k = \operatorname{sgn}\left[\left(\left(\boldsymbol{R} + \sigma^2\boldsymbol{A}^{-2}\right)^{-1}\boldsymbol{y}\right)_k\right] \tag{2-88}$$

由于考虑了背景噪声，因此最小均方误差 MMSE 检测器的误码性能优于解相关检测器，且不需要知道其他用户的特征码，没有增强噪声。其缺点是需要估计接收信号的幅度，对估计误差较敏感。

(3) 自适应多用户检测器。自适应多用户检测器利用了自适应滤波的原理，由于多径信道的时变性，使得本来已知的干扰用户扩频码结构信息参量变为时变的未知量，采用自适应方法可以直接找到这些滤波器参量，此外还可以避免复杂的计算，而且实现方便。自



适应多用户检测器分为非盲检测器和盲检测器，非盲检测器需要传输训练序列，而盲检测器则不需要。非盲检测器又分为自适应解相关检测器和自适应 MMSE 检测器。其中，自适应 MMSE 检测器又可划分为单用户型和多用户型两种。多用户自适应 MMSE 检测器结构如图 2.33 所示。

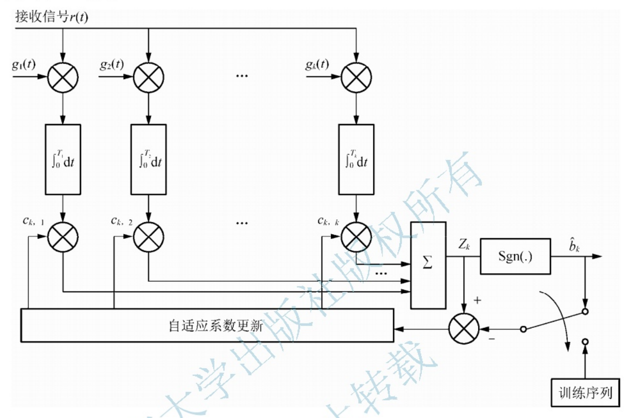


图 2.33　多用户自适应 MMSE 检测器

多用户自适应检测器采用了传统的匹配滤波器，不仅需要训练序列，而且需要知道其他用户的扩频序列。由于它的滤波器抽头只有 K 个，所以收敛速度很快。目前带训练序列的自适应算法已经比较成熟，主要包括随机梯度算法和递归最小二乘法 RLS 算法。在自适应算法中，人们最感兴趣的是仅仅需要知道一个发送波形或一个接收波形的盲自适应算法，盲自适应多用户检测器直接从业务信号本身提取信道状态信息，不需要其他用户信息和训练序列，只需要待测用户的观测数据，这样尤其是在快时变信道中，就不需要频繁发送训练序列，从而提高了系统的有效性和可靠性，由于它所需要的信息几乎与传统检测器相同，因此从本质上说是一种单用户抗多径自适应检测器。根据盲算法的代价函数不同，可分为以下几种：最小输出能量算法(MOE)、恒模盲检测算法(CMA)、基于子空间的盲自适应算法、还有 Griffiths 盲自适应算法、迫零解相关算法等等。

盲检测器的优点是无须训练序列，开销小、效率高、复杂度低，已成为当前的研究热点之一。但盲多用户检测技术的瓶颈是其收敛速度较慢，特别是在时变多径信道，这也是盲多用户检测器应用中需要解决的主要问题。

3) 非线性多用户检测技术

非线性多用户检测器中，主要有基于反馈判决的干扰抵消技术，它包括串行干扰抵消



检测器、并行干扰抵消检测器、串并混合型干扰抵消检测器以及迫零解相关检测器。另外，还有序列检测器、分组检测器、基于神经网络的多用户检测器和采用其他信号处理技术的检测器。下面主要介绍串行干扰抵消检测器和并行干扰抵消检测器。

(1) 串行干扰抵消检测器。串行干扰抵消检测器(SIC)(图 2.34)的基本思路是先解出各用户的数据比特，再根据信道估计值将该数据比特重构，即重构出期望用户的干扰信号，再从原接收的信号中减去这一重构的干扰信号，以达到干扰抵消的目的。

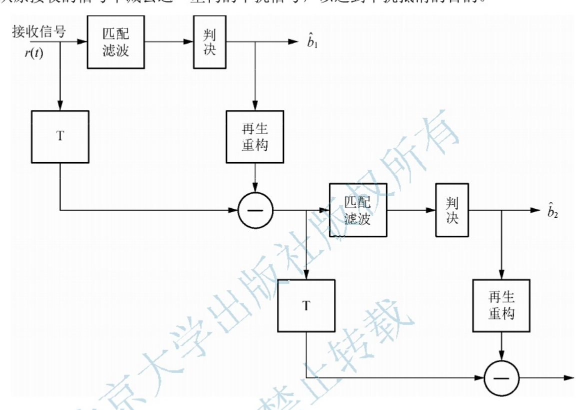


图 2.34　串行干扰抵消(SIC)检测器

串行干扰抵消法是消除多址干扰最简单、最直观的方法之一，其具体实现步骤如下。

步骤 1：根据接收到的各用户信号功率按强弱大小排队。

步骤 2：每次检测一个用户，先检测最强功率的用户，对信号进行硬判决。

步骤 3：根据判决结果、该信号的扩频码、估计出的幅度和相位信息得到该信号的估计值，从总的接收信号中减去重构的最强用户干扰估计值。

步骤 4：重复上述过程，再检测、重构和抵消次强干扰，依此类推。

串行干扰抵消检测器应用范围广泛，适用于同步和异步 CDMA 系统中。串行干扰抵消检测器的性能，很大程度上取决于用户接收信号的功率分布和初始数据估计的正确性和可靠性，功率分布差别越大，其性能越好，若初始比特估计出错，将导致干扰功率增强 4 倍。此外，串行干扰抵消检测器引入的处理时延较大，每个分组用户不能太多，一般取 4 个用户。

(2) 并行干扰抵消检测器。并行干扰抵消检测器由多级干扰抵消器构成，其中一级的结构如图 2.35 所示。并行干扰抵消检测器(PIC)的原理是根据接收信号的初始值或前级判决值，估计所有用户的干扰信号，然后对每个用户并行抵消掉所有干扰信号。

并行干扰抵消的步骤如下。

步骤 1：第 m 级经过匹配滤波器组得到的比特估计矢量，首先与估计的幅度信息相乘，然后重新扩频，生成重建的每个用户信息 $r_k(t)$。

步骤 2：部分求和单元对所有 K 路输入信号中的 $K-1$ 路求和，生成所有用户的多址干扰估计信号。

步骤 3：假设幅度和时延估计均理想，则接收信号减去多址干扰信号后的第 k 个用户信号为

$$y(t)-\sum_{i\neq k} r_i(t)=b_k A_k s_k(t-\tau_k)+n(t)+\sum_{i\neq k}(b_i-\hat{b}_i)A_i s_i(t-\tau_i) \tag{2-89}$$

步骤 4：第 m 级的判决信号送入下一级干扰抵消器，以生成更加可靠的数据估计信息。

并行干扰抵消检测器(PIC)与串行干扰抵消检测器(SIC)的设计思想是相似的，由于并行干扰抵消检测器是并行处理，克服了串行干扰抵消检测器延时大的缺点，而且无须在情况发生变化时进行重新排序，因此在各种多用户检测器中具有较高的实用价值。

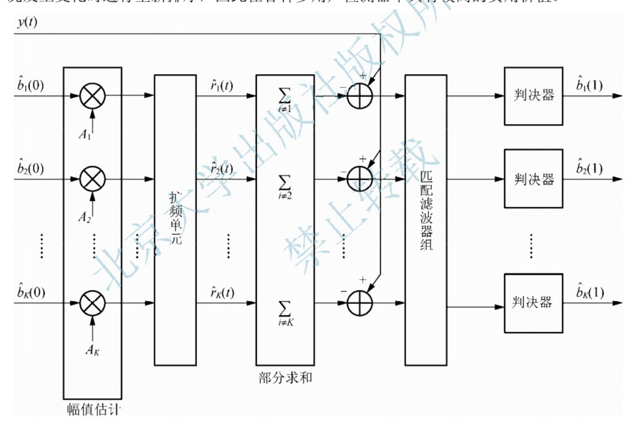


图 2.35　并行干扰抵消检测器

目前研究的多用户检测技术大部分是用在 DS-CDMA 系统中的，不论是同步 CDMA 系统还是异步 CDMA 系统都可应用多用户检测器。一般来说采用多用户检测器技术，必然涉及设备复杂度和体积等许多方面的问题，所以多用户检测器适于用在上行链路的基站中。但是随着多用户检测器技术研究的深入以及 DSP 和 IC 等技术的飞速发展，多用户检测器用在下行链路的移动台中，也是很有研究价值的。多用户检测还可应用于多载波 CDMA(MC-CDMA)系统中，该系统结合了正交频分复用(OFDM)技术与直接序列

CDMA(DS-CDMA)。在 MC-CDMA 系统中，由于多载波技术能够有效地克服衰落信道引起的码间干扰，因此可引入多用户检测，使之与多载技术相结合，研究频域与码域的二维信号处理技术，大大提高了系统的抗干扰能力，改善了系统性能。

2.5 信道容量

信道是信息传输的通道，其最大特征是信息的传输能力。信道容量则指信道中信息无差错传输的最大速率。在信息论中，信道分为离散信道和连续信道两大类，离散信道就是输入与输出信号都是取值离散的时间函数；连续信道是指输入和输出信号都是取值连续的。前者是广义信道中的编码信道，后者是调制信道。这里我们只讨论连续信道的信道容量。

1. 香农公式

连续信道的信道容量是由香农定理给出的，也称香农公式，具体描述为，在加性高斯白噪声的信道中，信道的信道容量为

$$C = B\log_2\left(1+\frac{S}{N}\right) \tag{2-90}$$

其中 B 为连续信道的带宽，单位为 Hz，S 为输入信号功率，N 为信道中加性高斯白噪声的功率，S/N 为信噪比，C 为信道容量，以信道每秒所能传输比特数为单位，记为 bps。

若噪声 $n(t)$ 的单边功率谱密度为 n_0，则在信道带宽 B 内的噪声功率 $N = n_0 B$。因此，可以得到香农公式的另一种表示形式

$$C = B\log_2\left(1+\frac{S}{n_0 B}\right) \tag{2-91}$$

香农公式表明若传输速率小于等于信道容量，则理论上可以实现无差错传输；若传输速率大于信道容量，则不可能实现无差错传输。

2. C 与 B、S、$N(n_0)$之间的关系

(1) 增大信号功率 S 可以增加信道容量，若信号功率趋于无穷大，则信道容量也趋于无穷大，即

$$\lim_{S\to\infty} C = \lim_{S\to\infty} B\log_2\left(1+\frac{S}{n_0 B}\right) \to \infty \tag{2-92}$$

(2) 减小噪声功率 N (或减小噪声功率谱密度 n_0)可以增加信道容量，若噪声功率趋于零(或噪声功率谱密度趋于零)，则信道容量趋于无穷大，即

$$\lim_{N\to 0} C = \lim_{N\to 0} B\log_2\left(1+\frac{S}{N}\right) \to \infty \tag{2-93}$$

(3) 增大信道带宽 B 可以增加信道容量，但不能使信道容量无限制增大。若信道带宽 B 趋于无穷大时，则信道容量的极限值为

$$\lim_{B\to\infty} C = \lim_{B\to\infty} B\log_2\left(1+\frac{S}{n_0 B}\right) = \frac{S}{n_0}\log_2 e \approx 1.44\frac{S}{n_0} \tag{2-94}$$

香农定理给出了通信系统所能达到的极限信息速率，这样的通信系统称为理想通信系统，但是，香农公式只证明了理想通信系统的存在性，却没有指出这种通信系统的实现方法。

【例 2.8】已知有线电话信道的带宽为 3.4kHz，试求输出信噪比为 30dB 时该信道的信道容量是多少？

解：根据香农公式可得该信道的信道容量

$$C = B\log_2\left(1+\frac{S}{N}\right)=3.4\times10^3\times\log_2(1+1000)=33.9(\text{Kbps})$$

注意：输出信噪比为 30dB，则 $S/N=1000$。

根据 C 与 B、S、$N(n_0)$ 之间的关系，对于一定的信道容量 C 来说，信道带宽 B、信噪比 S/N 及传输时间三者之间可以互相转换。若增加信道带宽，可以换来信噪比的降低；如果信噪比不变，那么增加信道带宽可以换取传输时间的减少等。

例如，若给定信道容量为 C，互换前的带宽和信噪比分别为 B_1 和 S_1/N_1，互换后的带宽和信噪比分别为 B_2 和 S_2/N_2，则有

$$C = B_1\log_2\left(1+\frac{S_1}{N_1}\right)=B_2\log_2\left(1+\frac{S_2}{N_2}\right)$$

上式也可写成

$$C = B_1\log_2\left(1+\frac{S_1}{n_0B_1}\right)=B_2\log_2\left(1+\frac{S_2}{n_0B_2}\right)$$

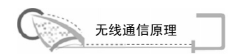
如果互换前信道带宽 $B_1=3\text{kHz}$，希望信息传输速率为 10^4 bps，则要求信噪比 $S_1/N_1\approx 9$。如果将带宽进行互换，设互换后的信道带宽 $B_2=10\text{kHz}$。这时，信息传输速率若仍为 10^4 bps，则所需要的信噪比为 $S_2/N_2=1$。可见，信道带宽 B 的变化可使输出信噪比也变化，而保持信息传输速率不变。这种信噪比和带宽的互换性在通信工程中有很大的用处。例如，在宇宙飞船与地面的通信中，飞船上的发射功率不可能做得很大，因此可用增大带宽的方法来换取对信噪比要求的降低。相反，如果信道频带比较紧张，如有线载波电话信道，这时主要考虑频带利用率，可用提高信号功率来增加信噪比，或采用多进制的方法来换取较窄的频带。前面我们讨论的是带宽和信噪比的互换。此外，带宽或信噪比与传输时间也存在着互换关系。


阅读材料

LiFi

电灯一直被视为发明家灵感闪现的象征，德国物理学家哈斯就由灯泡本身“点亮”了他的奇思妙想，他依赖一盏小小的灯，将看不见的网络信号变成了“看得见”的网络信号。哈斯和他在英国爱丁堡大学的团队最新发明了一种专利技术，利用闪烁的灯光来传输数字信息，这个过程被称为可见光通信(VLC)。相比 WiFi，可见光通信又被称为 LiFi，以示它可能给目前以 WiFi 为代表的无线网络传输技术带来革命性的改变。LiFi 使光信号传播在我们周围的环境中，利用光信号实现了数据的无线传输，只要自然光能到达的任何地方，就有 LiFi 的信号。

光和无线电波一样，都属于电磁波的一种，它们传送信号的基本原理是一致的。给普通的 LED 灯泡



装上微芯片，可以控制它每秒数百万次闪烁(灯亮表示 1，灯灭表示 0)。这样的变化能被光敏传感器捕获并解调出来，而人眼却不会察觉到 LED 灯的这种变化——因为它们闪烁得实在太快了。这样，二进制的数据就被快速编码成灯光信号并进行了有效的传输。LiFi 技术运用已铺设好的设备(无处不在的电灯)，只要在电灯上植入一个微小的芯片，电灯就变成了类似于 AP 的设备，使终端随时能接入网络。LiFi 的这一技术意味着，只要你拥有电灯，就可以获得无线互联网连接。灯光下的计算机，通过一套特制的接收装置，能够读懂灯光里的“莫尔斯密码”。研究者将网络信号接入一盏功率为 1W 的 LED 灯，与接收器相连的 4 台计算机就都能够上网，最高速率可达 3.25G，平均速率达到 150M，堪称世界最快的“灯光上网”。可以预想，未来任何有光的地方都可以成为潜在的 LiFi 数据传输源：在街头，利用路灯就可以下载电影；在家里，打开台灯就可以下载歌曲；在餐厅，坐在有灯光的地方就可以发微博；即便是在水下，只要有灯光照射就可以连接网络；汽车间依靠 LED 车灯来“对话”，飞机客舱里乘客利用头顶的 LED 阅读灯来连接网络…

与现有的 WiFi 相比，未来的可见光通信既安全又经济。WiFi 依赖看不见的无线电波传输，设备功率越来越大，局部电磁辐射势必增强，且无线信号穿墙而过，网络信息不安全。这些安全隐患，对 LiFi 来说可一扫而光，有灯光的地方，就有 LiFi 网络信号，关掉灯，网络全无。而且，可见光频谱的宽度无线电频谱的 1 万倍，这就意味着更大的带宽和更高的速率，网络设置又几乎不需要任何新的基础设施。目前 LiFi 还停留在实验室阶段，相信不远的将来，LiFi 也会像 WiFi 那样走进千家万户。

本 章 小 结

本章主要介绍了无线信道及其传输特性，无线通信系统的性能主要受到无线信道的制约，无线信道的传输特性直接影响到无线通信的质量。在无线通信的不同传播环境下，电磁波有两个方面的传输特性：大尺度衰落和小尺度衰落。大尺度衰落主要反映传播在宏观大范围内的空间距离上接收信号平均值的变化趋势，大尺度衰落主要有自由空间传播损耗、地面反射模型、散射和室外传播路径损耗模型。小尺度衰落反映了移动台在极小范围内移动时接收信号平均值的起伏变化趋势，小尺度衰落主要有频率选择性衰落、时间选择性衰落和空间选择性衰落，这些衰落特性分别与多径信号的时延扩展、多普勒扩展和角度扩展有关，这 3 种扩展又分别对应 3 组相关参数——相干带宽、相干时间和相干距离。衰落信道的统计特性主要介绍了瑞利分布和莱斯分布。

本章主要介绍了几种有效的抗信道衰落措施，包括：分集接收技术、RAKE 接收技术、均衡技术和多用户检测技术。分集接收技术包含分散接收和集中处理两重含义，分集方式主要有空间分集、时间分集、频率分集、极化分集、角度分集。RAKE 接收技术是一种典型的利用信号统计与信号处理技术将分集的作用隐含在被传输的信号之中的隐分集技术。均衡技术是接收端滤波器产生与信道特性相反的特性，用来减小或消除因信道的时变多径传播特性引起的码间干扰，主要是用它来实现对信道特性的均衡。多用户检测技术是一种抗多址干扰和多径干扰的抗衰落技术，它已成为第三代移动通信系统中宽带 CDMA 通信系统抗干扰的关键技术。

习　　题

2.1 填空题

(1) 电磁波的传播方式有__________、__________、__________和__________。

(2) 无线信道的小尺度衰落主要有__________、__________和__________。

(3) 在移动信道中市区的最大时延差为 5μs，室内的最大时延差为 0.04μs，这两种情况下的相关带宽分别为__________和__________。

(4) 分集接收的合并方式主要有__________、__________、__________和__________。

2.2 什么是大尺度衰落，其产生的原因是什么？什么是小尺度衰落，其产生的原因是什么？

2.3 室外传播路径损耗模型有哪些？各有什么特点？

2.4 常用的抗信道衰落的通信技术有哪些？

2.5 什么是分集接收技术？简述 RAKE 接收技术的原理。

2.6 简述均衡技术的原理。有哪几种均衡技术？各自的特点是什么？

2.7 简述盲均衡的原理。盲均衡技术的算法有哪些？

2.8 简述多用户检测技术的原理。有哪些多用户检测技术？

2.9 简述串行干扰抵消多用户检测和并行干扰抵消多用户检测的步骤。

2.10 某发射机通过天线发射出去的信号功率为 1W，工作频率为 2.4GHz，如果收发天线的增益均为 1.6，收发天线之间的距离为 1.6km，$L=1$。试求接收天线的接收功率和路径损耗。

2.11 GSM 系统的工作频率为 1.8GHz，基站天线高度为 50m，基站天线增益为 2，一个 GSM 移动台距离基站 5km，接收天线高度为 1.5m，接收天线增益为 1。试求：(1)自由空间传播路径损耗；(2)使用双线地面反射模型的情况下，计算传播路径损耗。

2.12 一个计算机用户打算买一个调制解调器，以便通过模拟电话线路发送数据。电话线路的信噪比为 25dB，话音频率的范围为 300～3400Hz，试求为保证无差错传输，该电话线路所能获得的最大信息传输速率。

第 3 章 无线通信编码技术

本章知识结构

- 无线通信编码技术
  - 信源编码技术
    - 信源编码原理
    - 语音编码
    - 数据压缩编码
    - 主要信源编码技术
      - 香农编码
      - 霍夫曼编码
      - 游程编码
      - 算术编码
      - 预测编码
      - 变换编码
        - 基本概念
        - 编译码原理
  - 信道编码技术
    - 差错控制方式
    - 信道编码原理
    - 主要信道编码技术
      - 线性分组码
        - 基本概念
        - 生成矩阵
        - 监督矩阵
        - 编译码原理
      - 循环码
        - 基本概念
        - 生成多项式
        - 生成矩阵
        - 编译码原理
      - 卷积码
      - Turbo码
      - LDPC码
      - 网格编码
        - 基本概念
        - 编译码原理



本章教学目的与要求

- 理解和掌握信源编码的概念和原理
- 了解语音编码和数据压缩编码的实现方法
- 理解和掌握香农编码、霍夫曼编码、算术编码、预测编码、变换编码的原理和实现方法
- 理解和掌握信源编码的概念和原理
- 了解信道编码的分类和差错控制方式
- 理解和掌握线性分组码的基本概念、监督矩阵和生成矩阵、线性分组码编译码原理
- 理解和掌握循环码的概念、生成多项式和生成矩阵、循环码的编译码原理和实现方法
- 理解和掌握卷积码、Turbo 码、LDPC 码、TCM 码等信道编码技术的概念和编译码原理



引言

信源和无线信道是无线通信系统的基本组成部分，信源发出的信息通过无线信道进行传输，要实现信源信息高速率、高质量、可靠地通过无线信道传送给信宿，需要解决两方面的问题：一是在不失真或允许一定失真的条件下，如何用尽可能少的符号传送信源信息，以便提高信息传输效率；二是在信道存在噪声等干扰的情况下，如何增强信号的抗干扰能力，同时又使得信息的传输速率最大。这两个方面的问题分别就是信源编码和信道编码的主要任务。本章将介绍无线通信中的主要编码技术。

【案例 3.1】商品条形码

我们进入超市和商场，各种各样的商品琳琅满目，仔细观察你会发现每一件商品在其包装上都会印有一组条形码，显然这些条形码是用来标识不同的商品，那你知道这组条形码的具体含义吗？

条形码是将宽度不等的多个黑条和空白，按照一定的编码规则排列，用以表达商品信息的图形标识符。标准商品条码由 13 位数码组成，可以分为 4 个部分，各段结构如下：

<u>×××</u>　<u>××××</u>　<u>×××××</u>　<u>×</u>

从左到右各段名称和内容分别如下。

(1) 前缀码：前 3 位，用于标识商品制造国家(或地区)，由国际上分配。690～695 都是中国的代码。如图 3.1 条码中“690”就代表中国。

(2) 制造商代码：第 4～7 位共 4 位，用于标识商品的生产厂商代码，由各厂商申请，国家主管机构分配。

(3) 商品代码：第 8～12 位共 5 位，用于标识具体产品类型，代表着厂内商品代码，由厂商自行确定。

(4) 校验码：第 13 位，依据一定的算法，由前面 12 位数码，通过计算而得到，用于校正在扫描过程中产生的识读错误。

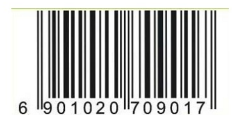


图 3.1　条形码示意图

【案例 3.2】二维码

二维码是用某种特定的几何图形按一定规律在平面(二维)上分布的黑白相间的图形记录数据符号信息。二维码在代码编制上巧妙地利用构成计算机内部逻辑基础的“0”、“1”比特流的概念，使用若干个与二进制相对应的几何形体来表示文字数值信息，通过图像输入设备或光电扫描设备自动识读以实现信息自动处理。

随着物联网的发展，二维码已经不是陌生的词汇，这个黑白小方格组成的矩阵图案，只需用手轻松一拍，就可获得意想不到的丰富信息。例如，轻轻扫一下贴有二维码的蔬菜，就可以知道它的产地、生长环境及过程；扫一下贴有二维码的水果，便知它的前世今生；在景区，用手机拍一下导览二维码，即可获得景点导览信息；刊登在海报上的打折信息，手机一拍，就可以拿到电子折扣券。二维码的应用还有很多，包括二维码购物、二维码查询、传情(文字、图片、视频、声音)、二维码寻宝、二维码看电影、二维码签到等(图 3.2)。在未来，二维码能做的更多，如匆忙上班的路上拿出手机拍个二维码，回到办公室前美味的早餐已等在桌上；下班回家，链接手机二维码，便能在家中试穿最新上市的时尚服饰；出外旅行不再需要导游，拍下二维码就可能穿越时空。



(a) 西瓜贴上二维码



(b) 羊肉贴上二维码



(c) 二维码名片

图 3.2　二维码应用

3.1　信 源 编 码

3.1.1　信源编码原理

为了提高数据信息传输的有效性而采取的编码称为信源编码，信源编码通过科学合理的对信源进行代码分配，将信源的原始符号按一定规则进行变换，以尽量减少数据传输长



度，压缩信源的冗余度，提高编码效率。信源编码分为无失真信源编码和限失真信源编码。无失真信源编码能够无失真地恢复原信源的数据信息，由于实际上完全无失真地传输信源信息是不可能的，而且实际中一般也不要求完全无失真地恢复信息，允许一定程度上的失真存在，所以实际使用的大都是限失真信源编码。

信源有离散信源和连续信源两大类，如文字、电报以及各类数据等都属于离散信源，而未经数字化的语音、图像则属于连续信源。离散信源编码可做到无失真编码，而连续信源编码则只能做到限失真编码。对连续信源进行编码时，可以通过对信源进行抽样(抽样定理)和量化转化为离散信源后再进行编码，所以信源编码主要是对离散信源编码，也就是将信源的原始符号序列，按一定的数学规则映射成码符号(码元)序列的过程，即信源编码的实质是对信源的原始符号按一定规则进行的一种变换。

信源编码的一般模型如图 3.3 所示，如果将信源编码器看作是一个网络，则它有 2 个输入和 1 个输出，分别是信源符号集合 $X=\{x_1,x_2,\cdots,x_N\}$、信道符号集合 A (也称为码表)和代码集合 $Y=\{y_1,y_2,\cdots,y_N\}$。信源编码就是将信源符号集中的符号 x_i 用信道的基本符号按照规定的编码方法进行编码并产生与输入序列一一对应的输出序列 y_i，这种输出的符号序列 y_i 称为码字，其长度 l_i 称为码字长度或简称码长。

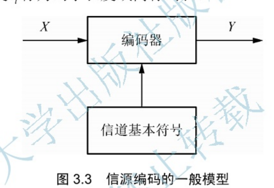


图 3.3　信源编码的一般模型

下面，给出一些码的定义。

二元码：是数字通信和计算机系统中最常用的一种码，码符号集为 $X=\{0,1\}$，所得码字都是二元序列。

等长码：码中所有码字的码长都相等。

变长码：码中所有码字的码长各不相同，任意码字由不同长度的码符号序列组成。

奇异码：码中存在有相同的码字(对应不同的信源符号)。

非奇异码：码中所有码字都不相同。

唯一可译码：由码的码字组成任意有限长的码字序列只能被唯一地译为对应的信源符号序列。

根据信源编码输出码长的特点，可以将信源编码分为等长编码和变长编码。等长编码对于信源不同的输出符号，码字的长度总是相同的，而变长编码产生的码长不完全相同。变长编码的基本思想是对于给定的信源，当信源符号不是等概率分布时，为了提高编码效率，给概率大的符号分配较短码字，概率小的符号分配较长码字，从而使得平均码长尽可能短。在序列长度 N 不是很大时，变长编码往往可实现高效的无失真信源编码。如表 3-1 中所示的“编码 1”为等长码，“编码 2”为变长码。

表 3-1　等长码与变长码

信源符号 x_i	符号出现概率 $p(x_i)$	编码 1	编码 2
x_1	$p(x_1)$	00	0
x_2	$p(x_2)$	01	01
x_3	$p(x_3)$	10	001
x_4	$p(x_4)$	11	101

根据香农理论，信源输出平均信息量定义为信源熵，即

$$\begin{aligned} H(X) &= E\left\{I\left[P(x_i)\right]\right\} = E\left[-\log_2 P(x_i)\right] \\ &= -\sum_{i=1}^{n} P(x_i)\log P(x_i) \end{aligned} \tag{3-1}$$

其中 n 为信源消息的可能种类数，$P(x_i)$ 为各种情况出现的概率。信源熵的含义是消息所含信息量的概率统计平均值。

对于等长编码，其编码效率为

$$\eta = \frac{H(X)}{l} \tag{3-2}$$

对于变长码，引入码的平均长度的概念作为衡量标准，设信源为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ P(x_1) & P(x_2) & \cdots & P(x_n) \end{bmatrix}$$

变长码编码产生的码字对应的长度分别为 l_1，l_2,…，l_n，则编码的平均长度为

$$\bar{L} = \sum_{i=1}^{n} P(x_i) l_i \tag{3-3}$$

变长编码时的编码效率为

$$\eta = \frac{H(X)}{\bar{L}} \tag{3-4}$$

3.1.2　语音编码与数据压缩编码

在第三代移动通信中，通信业务已不再是单一的语音业务类型，而是扩展到包含语音、数据和图像等在内的多媒体业务，因此无线通信中的信源编码将不仅包含语音编码、还包含各类图像压缩编码，多媒体数据压缩编码等方面的内容。

1. 语音编码

无线通信中的语音编码的基本要求是话音质量高、编码速率低；编译码延时短，延时应控制在 65ms 以内；抗干扰性能强；编译码器设计简单，功耗小，便于大规模集成化。语音编码的实现方法主要有波形编码、参数编码和混合编码三类。

波形编码是将时域的模拟语音波形信号经过采样、量化、编码后直接变换成数字码流的过程。波形编码的特点是编码速率较高，一般为 16～64Kbps，译码后质量较高、时延小。波形编码可以分为时域波形编码，如 PCM、AMPCM、ΔM 编码等，频域波形编码，如子



带编码(SBC)、自适应变换编码(ATC)等。波形编码在传统的公共电话交换网(PSTN)中广泛使用，由于波形编码的编码速率较高，无线传输时需要占用的频带较宽，所以不适合用于无线通信系统。

参数编码是从信源信号的某个域中提取特征参数，并变换成数字码流的过程。参数编码的特点是编码速率较低，一般为 1.2～4.8Kbps，占用带宽较小，较适合用于无线通信系统，译码后质量较低、时延较大。参量编码中话音质量只能达到中等水平，不能满足商用语音通信的要求。线性预测编码(LPC)就属于参量编码。

混合编码是波形编码和参数编码两者的混合，因此信号中既包含部分波形编码信息，又含有若干语音特征参量信息。混合编码技术综合了波形编码高语音质量和参量编码低编码速率的优点，从而使得混合编码以较低的编码速率获得较高的语音质量。混合编码方法是目前无线通信系统中广泛使用的语音编码技术。如欧洲 GSM 系统中采用的规则脉冲激励线性预测编码器(RPE-LTP)，美国 IS-54 标准的蜂窝系统使用的码率为 8Kbps 的矢量和激励线性预测(VSELP)语音编解码器，美国 IS-95 标准的蜂窝系统中采用的码激励线性预测编码(CELP)均属于混合编码。

2. 数据压缩编码

从 2.5G 开始到现在的第三代移动通信，通信业务不断丰富，除了语音业务之外，还包括多媒体等信息的传输，如文字、图像、视频等。在无线通信系统中这些多媒体所包含的数据量相对都比较大，如彩色图像、文件图像，特别是视频信息(电视、电影等)的数据量，在相同条件下要比语音的数据量大 1000 倍以上，那如何将这些数据量大的多媒体信息在有限的空间进行存储和传输呢？这就需要用到数据压缩编码技术了。

数据压缩编码是按照一定的编码机制对数据进行重新组织，力求用较少的数据来表示信息，以减少数据的冗余和存储的空间。数据压缩的可行性有两方面：一方面是视频图像或音频信号等原始信号源存在着很大的冗余度；另一方面是由于人的视觉特性对亮度信息很敏感，而对边缘的急剧变化不敏感，同时人的听觉也对部分频率的音频信号不敏感。因此将视频或音频数据进行压缩，再作解压处理后，人对恢复后的图像或音频信号仍有满意的主观感觉，这也说明人的感觉能接受这种数据压缩。

根据解码后数据与原始数据是否完全一致，数据压缩编码技术分为无损压缩和有损压缩两类。无损压缩是指解码的图像与原始图像严格相同，压缩比在 2∶1～5∶1，如霍夫曼编码、算术编码、游程编码等。有损压缩是还原的图像与原始图像存在一定的误差，但视觉效果一般可以接受，压缩比可以从几倍到上百倍，如 PCM(脉冲编码调制)、预测编码、变换编码以及新一代的数据压缩方法，如矢量量化和子带编码、基于模型的压缩、分形压缩和小波变换压缩等。

根据数据压缩方法的具体原理的不同，数据压缩编码技术分为以下几类。

(1) 预测编码。预测编码是根据空间中相邻数据存在相关性，利用前面一个或多个数据来预测下一个数据，然后对实际值和预测值的差(预测误差)进行编码。常用的预测编码方法有差分脉冲编码调制(DPCM)和自适应差分脉冲编码调制(ADPCM)等。

(2) 变换编码。变换编码是将数据信息由时域空间变换到频域空间上进行处理，在时

域空间上具有强相关性的数据信息，反映在频域某些特定的区域内能量常常被集中在一起，我们只需将主要注意力放在相对小的区域上，从而实现压缩。变换编码一般采用正交变换，如离散傅里叶变换(DFT)、离散余弦变换(DCT)、沃尔什哈达玛变换(WHT)和小波变换(WT)等来实现压缩算法。

(3) 统计编码。统计编码是根据信息熵原理，将出现概率大的符号用短码字表示，将出现概率小的符号用长码字表示。常用的统计编码方法有霍夫曼编码、香农编码以及算术编码。

(4) 子带编码。子带编码是将数据由时域变换到频域后，将其分割为若干的子频带，然后用不同的量化器进行量化，并分别进行编码。

(5) 模型编码。模型编码常用于图像编码中，编码时首先将图像中的边界、轮廓、纹理等结构特征找出来，然后保存这些参数信息，译码时根据结构和参数信息进行合成，恢复原始图像。具体的模型编码方法有轮廓编码、域分割编码、分析合成编码、识别合成编码、基于知识的编码和分形编码等。

下面介绍几种常用的信源编码技术。

3.1.3 香农编码

香农第一定理指出了信源与平均码长之间的关系，即$-\log_2 P(x_i) \leqslant l_i < -\log_2 P(x_i)+1$，同时也指出了可以通过编码使平均码长达到极限值。按照不等式$-\log_2 P(x_i) \leqslant l_i < -\log_2 P(x_i)+1$选择的码长所构成的码称为香农码，香农编码是采用信源符号的累积概率分布函数来分配码字的。

设某离散信源为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ P(x_1) & P(x_2) & \cdots & P(x_n) \end{bmatrix}$$

香农编码的步骤如下。

(1) 将各信源符号按概率递减的方式进行排列：$P(x_1) \geqslant P(x_2) \geqslant \cdots \geqslant P(x_n)$。

(2) 按香农不等式计算出每个信源符号的码长：$-\log_2 P(x_i) \leqslant l_i < -\log_2 P(x_i)+1$。

(3) 为了编成唯一可译码，计算信源第i个符号的累加概率$P_i = \sum_{k=1}^{i-1} P(x_k)$，$i = 2,3,\cdots,n$，$P_1 = 0$。

(4) 将累加概率P_i用二进制数表示。

(5) 取P_i对应二进制数的小数点后l_i位作为信源第i个符号的二进制码字。

【例 3.1】设有离散信源，按照概率大小排列后的概率分布如下，对这一信源进行香农编码，并计算平均码长及编码效率。

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ 0.2 & 0.19 & 0.18 & 0.17 & 0.15 & 0.10 & 0.01 \end{bmatrix}$$

解：香农编码的计算过程见表 3-2。



表 3-2　香农编码过程

信源符号 x_i	概率 $P(x_i)$	累加概率 $P_i=\sum_{k=1}^{i-1}P(x_k)$	$-\log_2 P(x_i)$	码长	P_i 对应的二进制数	码字
x_1	0.20	0	2.34	3	0	000
x_2	0.19	0.20	2.41	3	0.0011001	001
x_3	0.18	0.39	2.48	3	0.0110001	011
x_4	0.17	0.57	2.56	3	0.1001000	100
x_5	0.15	0.74	2.74	3	0.1011110	101
x_6	0.10	0.89	3.34	4	0.1110001	1110
x_7	0.01	0.99	6.66	7	0.1111110	1111110

此信源的熵为

$$H(X)=-\sum_{i=1}^{7}P(x_i)\log_2 P(x_i)=2.16(\text{bit/符号})$$

码的平均长度

$$\bar{L}=\sum_{i=1}^{7}P(x_i)l_i=3.14$$

编码效率

$$\eta=\frac{H(X)}{\bar{L}}=\frac{2.16}{3.14}=83.1\%$$

由例题可看出，香农编码所得的码字没有完全相同的，所以香农码是非奇异码。然而，由于香农码的编码效率不高，且冗余度较大，所以它不是最佳码，实用性也受到较大限制。



小知识

香农——信息论之父

克劳德·艾尔伍德·香农(1916—2001 年)，美国数学家，信息论及数字通信时代的奠基人。香农于 1916 年 4 月 30 日出生于美国密歇根州，据传，香农与伟大的发明家爱迪生有远亲关系。1936 年香农毕业于密歇根大学并获得数学和电子工程学士学位，1940 年获得麻省理工学院(MIT)数学博士学位和电子工程硕士学位。1941 年他加入贝尔实验室数学部，并一直工作到 1972 年，在此期间，1956 年香农成为麻省理工学院客座教授，并于 1958 年成为终生教授。香农的大部分时间是在贝尔实验室和麻省理工学院度过的。1948 年至 1949 年间，他先后发表了《通信的数学原理》和《噪声下的通信》，文章阐明了通信的基本问题，给出了通信系统的模型，提出了信息量的数学表达式，并解决了信道容量、信源统计特性、信源编码、信道编码等一系列基本技术问题。这两篇论文被视为信息论奠基之作，香农也因此一鸣惊人，被誉为“信息论之父”。

3.1.4 霍夫曼编码

霍夫曼编码是一种效率较高的变长、无失真信源编码方法。二进制霍夫曼编码的具体步骤如下。

(1) 将信源符号按概率从大到小的顺序排列。

(2) 给两个概率最小的信源符号各分配一个码位“0”和“1”，将这两个信源符号合并成一个新符号，并用这两个最小的概率之和作为新符号的概率，结果得到一个只包含 $(n-1)$ 个信源符号的新信源，这称为信源的第一次缩减信源，用 S_1 表示。

(3) 将缩减信源 S_1 的符号仍按概率从大到小的顺序排列，重复步骤(2)，得到只含 $(n-2)$ 个符号的缩减信源 S_2。

(4) 重复上述步骤，直至缩减信源只剩两个符号为止，此时所剩两个符号的概率之和必为 1。然后从最后一级缩减信源开始，依编码路径逆向返回，就得到各信源符号所对应的码字。

【例 3.2】对例 3.1 的信源进行二进制霍夫曼编码，并计算平均码长及编码效率。

解：霍夫曼编码的具体计算过程见表 3-3。

表 3-3　霍夫曼编码过程

信源符号 x_i	概率 $P(x_i)$	编码过程(缩减信源)						码字	码长
		S_1	S_2	S_3	S_4	S_5	S_1		
x_1	0.20	0.20	0.26	0.35	0.39	0.61	0 1.0	10	2
x_2	0.19	0.19	0.20	0.26	0.35	0 0.39	1	11	2
x_3	0.18	0.18	0.19	0.20	0 0.26	1		000	3
x_4	0.17	0.17	0.18	0 0.19	1			001	3
x_5	0.15	0.15	0 0.17	1				010	3
x_6	0.10	0 0.11	1					0110	4
x_7	0.01	1						0111	4

通过计算可得此信源的熵

$$H(X) = -\sum_{i=1}^{7} P(x_i)\log_2 P(x_i) = 2.16(\text{bit/符号})$$

码的平均长度

$$\bar{L} = \sum_{i=1}^{7} p(x_i) l_i = 2.72$$

编码效率

$$\eta = \frac{H(X)}{\bar{L}} = \frac{2.61}{2.72} = 96.3\%$$

由例 3.1 和例 3.2 可见，与香农编码相比，霍夫曼编码的平均码长更小，信息传输速率更大，编码效率更高。根据二进制霍夫曼编码可以推广到 m 进制霍夫曼编码，所不同的



只是每次把 m 个概率最小的符号分别用 $0,1,\cdots,m-1$ 等码元来表示，然后再将它们缩减为一个新的信源符号，其余步骤与二进制霍夫曼编码相同。

霍夫曼编码具有如下特点。

(1) 概率大的符号安排在离根节点较近的终端节点保证了概率大的符号对应于短码，概率小的符号对应于长码。

(2) 每次缩减信源的最长两个码字具有相同码长。

(3) 每次缩减信源的最后两个码字总是最后一位码元不同，前面各位码元相同。

(4) 编码不具有唯一性。

每次对信源缩减时，赋予信源最后两个概率最小的符号，分配码元“0”和“1”可以是任意的，即大概率符号或者合并后的符号集合分配的码元可以是“0”也可以是“1”，这种选择任意性导致得到不同的霍夫曼码，但这不会影响码字的长度；对信源进行缩减时，如果两个概率最小的符号合并后的概率与其他信源符号的概率相同，这两者在缩减信源中进行概率排序，其放置次序可以是任意的，故会得到不同的霍夫曼码。

香农编码与霍夫曼编码之间的区别：两者都考虑了信源的统计特性，使经常出现的信源符号对应较短的码字，进一步缩短信源的平均码长，从而实现了对信源的压缩；香农码编码结果唯一，但在很多情况下编码效率不是很高，霍夫曼码的编码方法不唯一；霍夫曼码对信源的统计特性没有特殊要求，编码效率比较高，对编码设备的要求也比较简单，因此综合性能优于香农码；霍夫曼码通常适用于多元信源，对于二元信源，必须采用合并符号的方法，才能得到较高的编码效率。

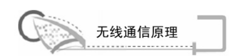
3.1.5　游程编码

香农码、霍夫曼码主要针对无记忆信源进行编码，当信源有记忆时上述编码的效率不高。游程编码是霍夫曼编码的改进和应用，它利用先后符号之间的关联性进行编码，编码效率较高，主要用于只有黑、白二值灰度的文件传真，如文件、报纸、表格、手写体字、图纸等。游程是指符号序列中各个符号连续重复出现而形成符号串的长度，又称游程长度或游长。游程编码就是将这种符号序列映射成游程长度和对应符号序列的位置的标志序列。如果知道了游程长度和对应符号序列的位置的标志序列，就可以完全恢复出原来的符号序列。

游程编码适用于对相关信源的编码，对二元相关信源，其输出序列往往会出现多个连续 0 或连续 1 的情况，连续出现的 0 符号称为“0”游程，连续出现的 1 符号称为“1”游程，对应连续同一符号的个数分别称为“0”游程长度和“1”游程长度。“0”游程和“1”游程总是交替出现，游程长度是随机的，其取值可以是 1，2，3，…，直至无穷，这样就将任何二元序列变换成了游程长度的序列(简称游程序列)，如二元序列 000101110010001，可变换成游程序列 3113213。

如果规定二元序列是以 0 开始的，那么第一个游程是“0”游程，第二个游程必为“1”游程，第三个游程又是“0”游程，以此类推。将任何二元序列变换成游程序列，这种变换是一一对应的，因此是可逆的、无失真的。

r 元序列同样也可以变换游程序列，在 r 元序列中，总共有 r 种游程，用 $L(i)$ 表示“i”



游程的长度，因此用 $L(i)$ 可以构成 r 元序列的游程序列。为了能保证一一对应的可逆变换性，这种变换必须再加一些分隔识别符号，才能区分游程序列中的某一个长度是 r 种游程中的哪一个长度；增加分隔识别符号把 r 元序列变换成游程序列再进行压缩编码通常效率不高，因此一般不对多元序列进行游程编码。

游程编码的具体方法：首先测定“0”游程长度和“1”游程长度的概率分布，以游程长度为元素，构造一个新的信源，对新的信源(游程长度序列)进行霍夫曼编码。因为游程长度是随机的、多值的，所以游程序列本身是多元序列，对游程序列可以按霍夫曼编码或其他编码方法进行处理以达到压缩码率的目的。因为游程长度的计数比较容易，得到游程长度后就可以从码表中找出码字输出。

若二元独立序列的统计特性已知，由二元独立序列与游程序列的一一对应性，可计算出游程长度序列的概率特性。设二元独立序列中符号 0 和 1 出现的概率分别为 p_0 和 p_1，则“0”游程长度 $L(0)$ 的概率为

$$p[L(0)] = p_0^{L(0)-1} p_1, \quad L(0) = 1, 2, 3, \cdots \tag{3-5}$$

同理可得“1”游程长度 $L(1)$ 的概率为

$$p[L(1)] = p_0 p_1^{L(1)-1}, \quad L(1) = 1, 2, 3, \cdots \tag{3-6}$$

且有 $\sum\limits_{L(0)=1}^{\infty} p[L(0)] = 1$，$\sum\limits_{L(1)=1}^{\infty} p[L(1)] = 1$。

“0”游程长度序列的熵为

$$H[L(0)] = \frac{H(p_0)}{p_1} \tag{3-7}$$

“1”游程长度序列的熵为

$$H[L(1)] = \frac{H(p_1)}{p_0} \tag{3-8}$$

“0”游程的平均长度 $\bar{l}_0 = 1/p_1$，“1”游程的平均长度 $\bar{l}_1 = 1/p_0$，假设“0”游程长度和“1”游程长度的霍夫曼编码效率分别为 η_0、η_1，则二元序列的游程编码效率为

$$\eta = \frac{H[L(0)] + H[L(1)]}{\dfrac{H[L(0)]}{\eta_0} + \dfrac{H[L(1)]}{\eta_1}} \tag{3-9}$$

根据式(3-9)，假设 $\eta_0 > \eta_1$，易知 $\eta_0 > \eta > \eta_1$。

当“0”游程和“1”游程的编码效率都很高时，采用游程编码的整体编码效率也很高，至少不会低于较小的那个游程的编码效率，要想游程的整体编码效率尽可能高，应尽量提高熵值较大的游程的编码效率，尽管游程长度可以从 1 一直到无穷长，但建立游程长度与霍夫曼码字之间的一一对应码表将是非常困难：游程越长，其出现的概率就越小，由霍夫曼码的编码规则，概率越小，码字越长；但小概率对应的长码字对平均码长影响很小，故对较长的二元序列，游程编码一般需采用截断处理。

3.1.6　算术编码

香农或霍夫曼等编码方法都是建立在信源符号与码字一一对应的基础上的，这种编码方法通常称为块码或分组码。使用分组码进行编码时，如果信源的符号较少，要取得好的编码效果只有采用扩展信源进行编码。算术编码跳出了分组编码的范畴，是一种非分组码的编码方法，它是一种从整个符号序列出发，采用递推形式进行编码的方法，信源符号和码字间不再是一一对应的关系。

算术编码的基本思路是：从整个符号序列出发，将各信源序列的概率映射到[0，1) 区间上，使每个符号序列对应于区间内的一点，也就是一个二进制的小数。这些点把[0，1) 区间分成许多小段，每段的长度等于某一信源序列的概率，在段内取一个二进制小数，其长度可与该序列的概率匹配，达到高效率编码的目的。这种方法与香农编码有些类似，只是它们考虑的信源对象有所不同，在香农编码中考虑的是单个信源符号，而在算术码中考虑的是信源符号序列。

如果信源符号集为 $A=\{a_1,a_2,\cdots,a_n\}$，信源序列 $\alpha=(a_{i1},a_{i2},\cdots,a_{il},\cdots,a_{iL}),a_{il}\in A$，则总共有 n^L 种可能的序列。由于考虑的是整个符号序列，因而整页纸上的信息也许就是一个序列，所以序列长度 L 一般都很大。在实际中很难得到对应信源序列的概率，一般我们从已知的信源符号概率 $P=\{p_1,p_2,\cdots,p_n\}$ 递推得到。

首先定义各符号的累积概率为 $P_i=\sum\limits_{l=1}^{i-1}p_l$，那么可得 $P_1=0$，$P_2=p_1$，$P_3=p_1+p_2=P_2+p_2$，…，$P_n=P_{n-1}+p_{n-1}$，由于 P_{i-1} 和 P_i 都是小于“1”的正数，可用[0，1) 区间内的两个点来表示，而 P_{i-1} 就是这两个点之间的长度，如图 3.4 所示。

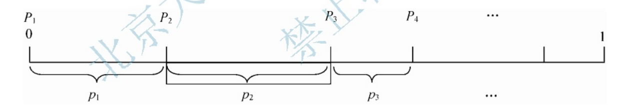


图 3.4　算术编码累积概率示意图

不同的信源符号有不同的概率区间，它们互不重叠，因此可以用这个小区间中的任意一点的取值，作为该信源符号的代码。我们所需确定的是这个代码所对应的长度，并使这个长度与信源符号的概率相匹配。对于整个信源符号序列而言，要把一个算术码字赋给它，则必须确定这个算术码字所对应的位于[0，1) 区间内的实数区间，即由整个信源符号序列的概率本身确定 0 和 1 之间的一个实数区间。随着符号序列中的符号数量的增加，用来代表它的区间会减小，而用来表达区间所需的信息比特数会增大。

每个符号序列中随着符号数量的增加，即信源符号的不断输入，用于代表符号序列概率的区间将随之减小，区间减小的过程如图 3.5 所示。区间宽度的递推公式为

$$A(\boldsymbol{\alpha}r)=A(\boldsymbol{\alpha})p(r)=p(\boldsymbol{\alpha})p(r)=p(\boldsymbol{\alpha}r),\quad r=0,1 \tag{3-10}$$

累积概率的递推公式为

$$P(\boldsymbol{\alpha} r)=P(\boldsymbol{\alpha})+p(\boldsymbol{\alpha})P(r),\ \ r=0,1 \tag{3-11}$$

其中 $P(\boldsymbol{\alpha})$ 为信源符号序列的累积概率， $p(\boldsymbol{\alpha})$ 为信源符号序列的联合概率，而 $P(0)=0$ ，$P(1)=p(0)$ 。

我们可取该小区间内的一点来代表这个信源符号序列，那么选取此点的方法可以有多种，实际中常取小区间的下界值。对信源符号序列的编码方法也可有多种，下面介绍常用的一种算术编码方法。

将信源符号序列 $\boldsymbol{\alpha}$ 的累积概率值写成二进位的小数 $P(\boldsymbol{\alpha})=0.c_1c_2\cdots c_L$， $c_i\in\{0,1\}$ ，取小数点后 L 位，若后面有尾数，就进位到第 L 位，并使 L 满足

$$L=\left\lceil \log_2\frac{1}{p(\boldsymbol{\alpha})}\right\rceil \tag{3-12}$$

式中 $\lceil x \rceil$ 表示大于或等于 x 的最小整数，这样就得到信源符号序列所对应的一个算术码 $c_1c_2\cdots c_L$ 。

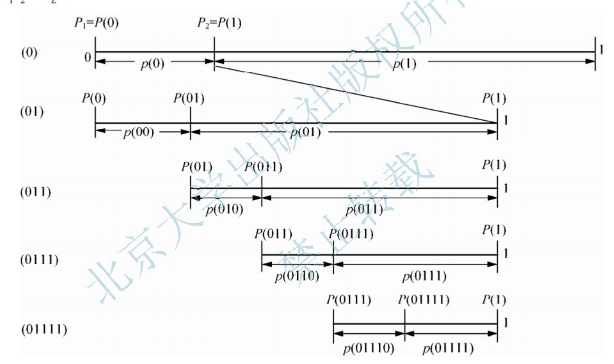


图 3.5　算术编码区间减小的过程示意图

【例 3.3】设二元无记忆信源 $X=\{0,\ 1\}$ ，其中 $p(0)=\dfrac{1}{4}$ ， $p(1)=\dfrac{3}{4}$ 。对二元序列 $\boldsymbol{\alpha}=11111100$ 进行算术编码。

解： 根据算术编码方法，先计算信源符号序列 $\boldsymbol{\alpha}$ 的联合概率

$$p(\boldsymbol{\alpha})=p(11111100)=p(1)^6\,p(0)^2=(3/4)^6(1/4)^2=0.01112366$$

信源符号序列的算术码字长度为

$$L=\left\lceil \log_2\frac{1}{p(\boldsymbol{\alpha})}\right\rceil=\lceil 6.49\rceil=7$$



再计算信源符号序列的累积概率，按递推公式有

$$
\begin{aligned}
P(\boldsymbol{\alpha}) = P(11111100) &= p(0) + p(10) + p(110) + p(1110) + p(11110) + p(111110) \\
&= (1/4) + (3/4)(1/4) + (3/4)^2(1/4) + (3/4)^3(1/4) + (3/4)^4(1/4) + (3/4)^5(1/4) \\
&= 0.82202148 \\
&= (0.1101001001\cdots)_2
\end{aligned}
$$

累积概率值 $P(\boldsymbol{\alpha})$ 即为输入符号序列“11111100”区间的下界值。取 $P(\boldsymbol{\alpha})$ 二进制表示小数点后 L 位，得到信源符号序列的算术码字为 1101010。对输入信源符号序列进行算术编码后的平均码长 $\bar{L} = \frac{7}{8}$，编码效率 $\eta = \frac{H(X)}{\bar{L}} = \frac{0.811}{7/8} = 92.7\%$。

因为这里需编码的序列长为 8 位，所以一共要将半开区间[0, 1) 分成 256 个小区间，以对应任一个可能的序列。由于任意一个码字必在某个特定的区间，所以其解码具有唯一性。

从性能上来看，算术编码具有许多优点，它所需的参数较少、编码效率高、编译码简单，在实际实现时，常用自适应算术编码对输入的信源序列自适应地估计其概率分布。算术编码在图像数据压缩标准(如 JPEG)中应用广泛。



小知识

JPEG

在计算机处理中，JPEG 是一种广泛适用的图像压缩标准方式。JPEG 是 Joint Photographic Experts Group(联合图像专家组)的首字母缩写，采用这种压缩格式的文件一般就称为 JPEG，此类文件的扩展名为 .jpg 或 .jpeg，其中在主流平台最常见的是. jpg 后缀。JPEG 格式是一种支持 8 位和 24 位色彩的压缩位图格式，它可以把文件压缩到最小的格式，在 Photoshop 软件中以 JPEG 格式储存时，提供 11 级压缩级别。JPEG 格式的应用非常广泛，特别是在网络和光盘读物上，都能找到它的身影，各类浏览器均支持 JPEG 这种图像格式，由于 JPEG 格式的文件尺寸较小，下载速度快，是目前网络上最流行的图像格式。

3.1.7　预测编码

预测编码是数据压缩三大经典技术(统计编码、预测编码、变换编码)之一，它是建立在信源数据相关性的基础上的。由信息理论可知，对于相关性很强的信源，条件熵可远小于无条件熵，因此人们常采用尽量解除相关性的办法，使信源输出转化为独立序列，以利于进一步压缩码率。

预测编码的基本思想不是直接对信源符号进行编码，而是通过充分利用信源符号之间的统计相关性，先将信源输出符号序列进行预测变换，再对信源输出与被预测值的差值进行编码，这里正是由于信源符号间存在相关性，所以才使预测成为可能，对于独立信源，预测就没有可能。预测的理论基础主要是估计理论，估计就是用实验数据组成一个统计量作为某一物理量的估值或预测值。若估值的数学期望等于原来的物理量，就称这种估计为无偏估计；若估值与原物理量之间的均方误差最小，就称之为最佳估计，基于这种方法进行预测，就称为最小均方误差预测，这种预测也是最佳预测，要实现最佳预测就是要找到计算预测值的预测函数。预测编码的基本原理图如图 3.6 所示，编码端主要由一个符号编码器和一个预测器组成，译码端主要由一个符号译码器和一个预测器组成。

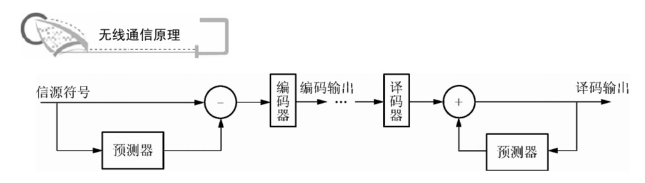


图 3.6　预测编码器

如果信源的相关性很强，则采用实际值与预测值的差值进行预测编码可获得较高的压缩率。由于相关性很强的信源可较精确地预测待编码的值，使得这个差值的方差将远小于原来的信源取值，所以在同样失真要求下，量化级数可大大减少，从而较显著地压缩码率。

预测编码的特点：在差值预测编码中所能取得的压缩率与预测误差序列所产生的熵的减少量直接有关；通过预测可消除相关，所以预测误差的概率分布一般在零点附近有一个高峰，并且与输入信源分布相比其方差较小；误码易于扩散，抗干扰能力强。

3.1.8　变换编码

在许多情况下，信源输出符号之间具有很强的相关性，如果按照前面介绍的霍夫曼、算术编码等编码算法来实现高效数据压缩，则需要使用扩展信源进行编码，因此就需要知道序列之间的联合概率分布。由于编码算法本身就比较复杂，因此当信源符号数量较多时，如果采用较长的序列进行编码，就会进一步增加编码系统的复杂度。为了提高编码效率，同时降低编码复杂度，可以对信源输出的数据进行变换，以去除数据之间的相关性，即采用变换编码方式，将数据之间的相关冗余变换为系数之间的统计冗余，从而降低系统编码复杂度，提高编码效率。

变换编码的基本原理是将信号从一种信号空间变换到另一种有利于压缩编码的信号空间(如信号由时域到频域的变换)，然后进行编码。信号经过变换后，变换系数之间的相关性常常显著下降，并常有能量集中于低频或低系数区域的特点，这样就容易实现码率的压缩，并还可大大降低数据压缩的难度。一般的变换编码系统框图如图 3.7 所示。变换编码常常用于图像的压缩编码，图像变换域编码通常将空间域相关的像素点通过某种变换关系映射到另一个频域上，使变换后的系数之间的相关性降低，在变换后的频域上满足所有系数相互独立，能量集中于少数几个系数上，且这些系数集中在一个最小的区域内。然后对这些变换系数进行量化、编码、传输，这样对变换系数进行压缩编码，往往比直接对图像数据本身进行压缩更容易获得高的效率。

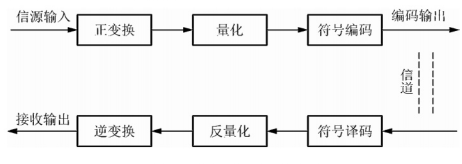


图 3.7　一般的变换编码系统框图





在实际编码系统中，变换编码一般采用正交变换，常用的变换方法有傅里叶变换(DFT)、离散余弦变换(DCT)、沃尔什变换(WHT)、哈尔变换、哈达玛变换和小波变换等等，不同的变换会有不同的图像压缩效果，在语音、图像编码等应用中，由于离散余弦变换算法简单、有效，因此得到了广泛的应用。

对于变换编码具体实现的各种变换方法在这里就不详细讲述了，如果有需要可以查阅相关参考文献。

3.2　信 道 编 码

3.2.1　差错控制方式

无线信道中要传送的电磁波信号，就像要运送的货物一样，半路上随时可能碰到各种各样的状况，我们知道要实现货物的安全运送，就要将运送的货物进行巧妙的伪装，以防货物被抢劫或丢失，那么要实现电磁波信号的可靠传输，同样要采取类似的手段，而信道编码就是提高信息传输可靠性的有效手段。信道编码是在信源编码后的数字序列上增加一些冗余信息，增加信息传输的抗干扰能力，以抵抗各种噪声和衰落的影响。在实际信道中传输数字信号时，由于信道特性不理想及加性噪声的影响，接收端所收到的数字信号不可避免地会产生错码，影响通信质量，信道编码通过加入冗余码来减少误码，增加了信息的冗余度，通常冗余度是特定的、有规律的，故可利用其在接收端进行检错和纠错。因此信道编码也称为差错控制编码、抗干扰编码或纠错编码。

常用的差错控制方式主要包括前向纠错(FEC)、检错重发(ARQ)、混合纠错(HEC)和信息反馈(IRQ)，这 4 种差错控制方式如图 3.8 所示。

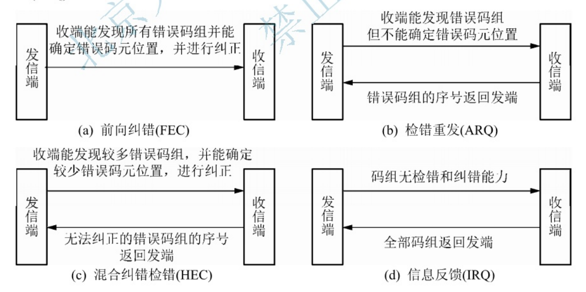


图 3.8　差错控制方式

1. 前向纠错

在前向纠错(FEC)方式中，发信端将信息码经信道编码后变成能够纠正错误的码，然

后通过信道发送出去；收信端收到这些码组后，根据与发信端约定好的编码规则，通过译码能自动发现并纠正因传输带来的数据错误。前向纠错方式的优点是不需要反馈信道，适合于单向信道的场合，同时也适合于一点发送多点接收的广播方式。因为不需要对发信端反馈信息，所以接收信号的延时小、译码实时性好。这种纠错系统的缺点是设备复杂、成本高，且纠错能力越强，编译码设备就越复杂。

2. 检错重发

在检错重发(ARQ)方式中，发信端将信息码编成能够检错的码组发送到信道，收信端收到一个码组后进行检验，将检验结果(有无误码)通过反向信道反馈给发信端作为对发信端的一个应答信号(ACK)。发信端根据收到的应答信号(ACK)做出是继续发送新的数据还是把出错的数据重新发送的判断。检错重发系统又可分为停发等候重发、返回重发和选择重发 3 种工作方式。

停发等候重发系统中，发信端在某一时刻向收信端发送一个码组，收信端收到该码组后经检验若未发现错误，则发送一个应答信号(ACK)给发信端，发信端收到 ACK 信号后再发送下一个码组；如果接收端检验出有错误，则发送一个否认信号(NAK)，发信端收到 NCK 信号后重发前一个码组，并在此等待发送端的 ACK 或 NAK 信号。从上述过程可见，由于发信端要等待收信端的应答信号，所以发送过程是间歇式的，因此数据传输效率不高，但由于该系统原理简单，在计算机通信中仍然得到应用。

返回重发系统中，发信端连续不断地发送一个又一个码组，不再等候 ACK 信号，如果收信端发现错误并发回 NAK 信号，则发信端从下一个码组开始重发前一段 N 个码组，N 的大小取决于信号传输和处理所造成的延时，也就是发信端从发错误码组开始，到收到 NAK 信号为止所发出的码组个数。这种返回重发系统的传输效率比停发等候系统有很大改进，在许多数据传输系统中应用广泛。

选择重发系统中，发信端也是连续不断地发送一个又一个码组，收信端检测到错误后发回 NAK 信号，与返回重发系统不同之处是，发信端不是重发前 N 个码组，而是只重发有错误的那一组。显然，选择重发系统传输效率最高，但控制最为复杂，价格也最贵，在收、发两端都要求有数据缓存器。

3. 混合纠错

混合纠错(HEC)方式是前向纠错和检错重发两种方式的结合。发信端发送的不仅是能够检测错误的码，而且对超出纠错能力的错误还具有一定的检测能力。收信端收到码组后，检验差错情况，如果错码在码的纠错能力范围内，则自动纠错；如果超出了码的纠错能力，则能检测出来，然后经反馈信道请求发信端重发。混合纠错方式在实时性和译码复杂性方面是前向纠错和检错重发方式的折中，可达到较低的误码率，较适合于环路延迟大的高速数据传输系统，近年来，混合纠错方式在数据通信系统中应用也较广泛。

4. 信息反馈

在信息反馈(IRQ)方式中，收信端将收到的信息全部由反向信道送回发信端，发信端通过比较发送的信息与送回的信息来检验是否发生错误，并将认为错误的信息的原信息再



次发送，直到发信端没有发现错误为止。信息反馈方式的优点是不需要纠错、检错的编译器，设备简单，其缺点是需要和前向信道相同的反向信道，实时性差，环路时延较大，发信端需要有一定容量的存储器来存储发送的码组，数据传输速率越高，所需存储容量就越大。

3.2.2　信道编码的分类

信道编码有多种实现方式，主要的分类如下。

按照信道编码的不同功能可分为：检错码和纠错码。检错码仅能检测误码，例如，在计算机串行通信中常用到的奇偶校验码等；纠错码具有检错能力的同时也可以纠正误码，当发现不可纠正的错误时可以发出出错指示。

按照信息码元和监督码元之间的检验关系可分为：线性码和非线性码。若信息码元与监督码元之间的关系为线性关系，则称为线性码；反之，若两者不存在线性关系，则为非线性码。

按照信息码元和监督码元之间的约束方式不同可分为：分组码和卷积码。在分组码中，编码后的码元序列每 n 位分为一组，其中 k 个信息码元，r 个监督位，$r=n-k$，监督码元仅与本码字的信息码元有关；而在卷积码中，码组中的监督码元不但与本组信息码元有关，还与前面码组的信息码元也有约束关系，就像链条那样一环扣一环，因此卷积码也称连环码或链码。

按照信息码元在编码后是否保持原来的形式可分为：系统码和非系统码。在系统码中，编码后的信息码元保持原样不变，非系统码中的信息码元发生变化。除了个别情况，系统码的性能大体上与非系统码相同，但是非系统码的译码较为复杂，因此，系统码得到了广泛的应用。

按照纠正错误的类型不同可分为：纠正随机错误码和纠正突发错误码。纠正随机错误码主要用于发生零星独立错误的信道；纠正突发错误码用于以突发错误(指成串出现的错误)为主的信道。

按照信道编码所采用的数学方法不同可分为：代数码、几何码和算术码。其中代数码是目前发展最为完善的编码，线性码就是代数码的一个重要的分支。

3.2.3　信道编码基本原理

信道编码的基本思想是在被传送的信息中加入一些监督码元，在收和发之间建立某种检验关系，为了说明信道编码的基本原理，首先介绍一些相关概念。

码长：码组中码元的数目。

码重：码组中非零码元的个数。对于二进制码，码重是码组中 1 的数目。如：码组 000 的码重为 0，码组 10100，码长为 5，码重为 2。

码距：两个长度相同的不同码组中对应位不同的码元数目。将二维空间距离称为汉明距离，三维、四维等空间的几何距离称为欧氏距离。如码组 11001100 和 10010111 之间的码距为 5。

最小码距：在码组集合中全体码组之间距离的最小值，记为 $d_{\min}$。对于二进制码组而

言，两个码组之间的模 2 相加，其不同的对应位必为 1，相同的对应位必为 0，因此，两个码组之间模 2 相加得到的码重就是这两个码组之间的汉明距离，码组之间的最小距离是衡量该码组检错和纠错能力的重要依据，也是信道编码的重要参数，$d_{\min}$ 越大，则从一个码字错为另一个码字的可能性越小，其检错和纠错能力也就越强。

信道编码的抗干扰能力与最小码距之间有以下关系。

(1) 如果码组中码字间的最小码距满足

$$d_{\min} \geqslant e+1 \tag{3-13}$$

则该码组中的码字具有检测 e 个错码的能力。当码字之间的最小距离 $d_{\min}$ 比错码个数 e 至少还要大 1 时，只要码字中出现错码的个数不超过 e，则接收端都能发现这样的差错。这个关系的几何解释如图 3.9(a)所示。其中 C_1 和 C_2 分别表示两个码距为 d_0 的码字，若 C_1 发生 e 个错误，则 C_1 就变成以 C_1 为球心、e 为半径的球面上的码字，为了能将这些码字分辨出来，它们必须与其最近的码字 C_2 有一位的距离差别，即 C_1 和 C_2 之间的最小距离为 $d_{\min} \geqslant e+1$。

(2) 如果码组中码字间的最小码距满足

$$d_{\min} \geqslant 2t+1 \tag{3-14}$$

则该码组中的码字具有纠正 t 个错码的能力。因为当码字中某个准用码字发生 t 个错码后，形成的错误码字只要与另外任一准用码字错 t 个码之后的码距至少为 1，则这两个码字就变化混淆。这个关系的几何解释如图 3.9(b)所示。C_1 和 C_2 分别表示两个码距为 d_0 的码字，若 C_1 发生 t 个错误，则 C_1 就变成以 C_1 为球心、t 为半径的球面上的码字，若 C_2 发生 t 个错误，则 C_2 就变成以 C_2 为球心、t 为半径的球面上的码字，为了在出现 t 个错误之后，仍能分辨出 C_1 和 C_2，那么，C_1 和 C_2 之间的距离应大于 $2t$，最小距离也应当使两个球体表面相距为 1，即满足 $d_{\min} \geqslant 2t+1$。

(3) 如果码组中码字间的最小码距满足

$$d_{\min} \geqslant t+e+1 \tag{3-15}$$

则该码组中的码字具有纠正 t 个错码、检测 e 个错码的能力。若接收码字与某一准用码字之间的距离在纠错能力范围内则对其进行纠错，若接收码字与另外任一准用码字之间的距离都超过 t，则只能进行检错，检错能力范围为 e。显然，某一准用码字出现 e 个错码后的错误码字与另外任一准用码字之间的距离至少应为 $t+1$，否则就会进入另一个准用码字的检错能力范围而被错纠为另一个码字。这个关系的几何解释如图 3.9(c)所示。C_1 和 C_2 分别表示两个码距为 d_0 的码字，当码字出现 t 个或小于 t 个错误时，系统按照纠错方式工作；当码字出现大于 t 个而小于 e 个错误时，系统按照检错方式工作；若 C_1 发生 t 个错误，C_2 发生 t 个错误时，既要纠正 C_1 的错又要检测 C_2 的错，则 C_1 和 C_2 之间的距离应大于 $t+e$，即满足不等式 $d_{\min} \geqslant t+e+1$。

以上 3 种关系是编码理论中的重要基本关系式，对任意一类编码都适用。衡量信道编码的性能常用编码效率和编码增益来描述。

(1) 编码效率。码字中信息位数为 k，监督位数为 r，码长 $n=k+r$，定义编码效率为信息码位数 k 在码字 n 中所占的比重，表达式为

$$R_{\mathrm{c}} = \frac{k}{n} = \frac{n-r}{n} = 1-\frac{r}{n} \tag{3-16}$$



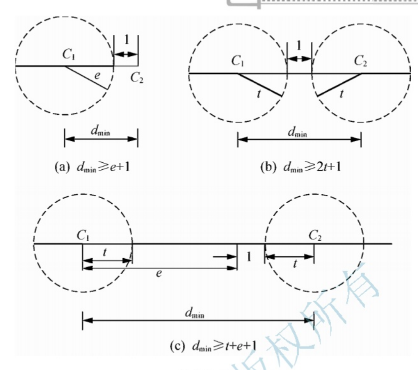


图 3.9　纠错能力的几何解释

在信道编码过程中，监督位越多纠错能力就越强，但编码效率就越低。R_c 表示传输信息的有效性，它说明了信道的利用效率，是衡量信道编码性能的主要指标。

(2) 编码增益。在数字通信中，信噪比通常用 E_b/n_0 来表示，其中 E_b 为信号的比特能量，n_0 为噪声功率谱密度(单边)，编码增益为

$$G_c = \left(\frac{E_b}{n_0}\right)_{UC} - \left(\frac{E_b}{n_0}\right)_C (\mathrm{dB}) \tag{3-17}$$

其中 $\left(\frac{E_b}{n_0}\right)_{UC}$ 和 $\left(\frac{E_b}{n_0}\right)_C$ 分别表示未编码和编码后所需要的信噪比。

3.2.4　常用的简单信道编码

1. 恒比码

恒比码也称等重码，恒比码的每个码组中 1 的个数与 0 的个数恒定，在接收端只要发现接收码元的“0”和“1”的数目比值发生变化，则可判断是否产生误码。恒比码的检错能力较强，除对换差错(1 和 0 成对的产生错误)不能发现外，其他各种错误均能发现。若码长为 n，码重为 w，则恒比码的许用码字个数为 C_n^w，禁用码字个数为 $2^n - C_n^w$。目前我国电传通信中普遍采用 3:2 码，该码共有 $C_5^3 = 10$ 个许用码字，用来传送 10 个阿拉伯数字，见表 3-4。传输汉字码时，每个汉字用 4 个阿拉伯数字表示，而每个阿拉伯数字又用 5 位二进制符号构成的码组表示，每个码组的长度为 5，其中恒有 3 个 1，因此，这种码又称 5 中取 3 恒比码。

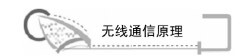


表 3-4　我国电传采用的 5 中取 3 恒比码

阿拉伯数字	编　　码	阿拉伯数字	编　　码
0	01101	5	00111
1	01011	6	10101
2	11001	7	11100
3	10110	8	01110
4	11010	9	10011

国际无线电报通信系统中，采用 3∶4 码即 7 中取 3 码，码组中规定总是有 3 个 1，因此，这种码共有 $C_7^3 = 35$ 个许用码字，分别表示 26 个英文字母及其他符号，见表 3-5。实践证明，采用这种码的国际电报通信的误码率保持在10^{-6}以下。恒比码的主要优点是简单，而且适于用来传输电传机或其他键盘设备产生的字母和符号。

表 3-5　国际通用的 7 中取 3 恒比码

字　　符		码	字　　符		码
A	−	0011010	S	′	0101010
B	?	0011001	T	5	1000101
C	:	1001100	U	7	0110010
D	+	0011100	V	=	1001001
E	3	0111000	W	2	0100101
F	%	0010011	X	/	0010110
G		1100001	Y	6	0010101
H		1010010	Z	+	0110001
I	8	111000	回行		1000011
J		0100011	换行		1011000
K	(	0001011	字母键		0100110
L	)	1100010	数字键		0001110
M	.	1010001	间隔		1101000
N	,	1010100	(不用)		0000111
O	9	1000110	RQ		0110100
P	0	1001010	α		0101001
Q	1	0001101	β		0101100
R	4	1100100			

2. 奇偶监督码

奇偶监督码是一种最简单的检错码，也称奇偶校验码，奇偶校验码有奇校验码和偶校验码两种，其编码规则是将信息码先分组，形成多个许用码组，在每一个许用码组最后(最



低位)加上一位监督码元即可。加上监督码元后使该码组中 1 的个数为奇数的编码称为奇校验码，为偶数的编码称为偶校验码。假设奇偶监督码的码组长度为 n，码字表示为 $a_{n-1},a_{n-2},\cdots,a_0$，前 $n-1$ 位为信息位，最后一位 a_0 为监督位，则奇校验关系式为

$$a_{n-1}\oplus a_{n-2}\oplus\cdots\oplus a_0=1 \tag{3-18}$$

其含义是奇校验监督位的选择要保证信息位与附加位模 2 和为 1。

偶校验关系式为

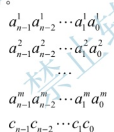
$$a_{n-1}\oplus a_{n-2}\oplus\cdots\oplus a_0=0 \tag{3-19}$$

其含义是偶校验监督位的选择要保证信息位与附加位模 2 和为 0。

根据奇偶校验的规则可以知道，当码组中的误码为偶数时，校验失效。如有两位发生错误，会有这样几种情况：00 变成 11、11 变成 00、01 变成 10、10 变成 01，这时奇校验码中 1 的个数仍为奇数，偶校验码中 1 的个数仍为偶数，可见无论哪种情况出现都不会改变码组的奇偶性。因此，奇偶校验码的监督关系式只能判断码元中存在奇数个错误的情况，对偶数个错误不能识别，并且只能检错，不能纠错。

3. 二维奇偶监督码

二维奇偶监督码是针对奇偶监督码检错能力不高、特别是不能检测突发错误(成串出现的错误)的缺点而提出的。将经过奇偶校验编码的码元序列排成方阵，每一个码组写成一行，然后再按列的方向增加第二位监督位就构成了二维奇偶监督码，如图 3.10 所示。图中 $a_0^1a_0^2\cdots a_0^m$ 为 m 行奇偶监督码中的 m 个监督位，$c_{n-1}c_{n-2}\cdots c_0$ 为按列进行第二次编码所增加的监督位，它们构成了一个监督位行。



$$\begin{gathered}a_{n-1}^1a_{n-2}^1\cdots a_1^1a_0^1\\a_{n-1}^2a_{n-2}^2\cdots a_1^2a_0^2\\\cdots\\a_{n-1}^ma_{n-2}^m\cdots a_1^ma_0^m\\c_{n-1}c_{n-2}\cdots c_1c_0\end{gathered}$$

图 3.10　二维奇偶监督码

二维奇偶校验码优于一维奇偶校验码，主要在于二维奇偶校验码可以检测奇数个错误，也可以检测出某些偶数个错误。当一组码元发生偶数个错误时，虽然行监督位不能检测出来，但列监督位可以检测出来，但要注意的是，当错码构成矩形区域时，二维奇偶校验码也不能检测出来。二维奇偶校验码不仅可以用来检错，也可以纠错，例如，当码组中仅在一行中有奇数个错误时，则能够确定错误的位置，从而将其纠正。

4. 正反码

正反码中监督位数目与信息位数目相同，当信息位中“1”的个数为奇数时，监督位与信息位一致，当信息位中“1”的个数为偶数时，监督位是信息位的反码。例如，若信息位为 10011，则正反码组为 1001110011；若信息位为 10001，则正反码组为 1000101110。正反码是一种简单的纠错编码，长度为 10 的正反码，可以纠正 1 位错码。

3.3 线性分组码

3.3.1 线性分组码的基本概念

将信源输出的信息序列每k个码元相继分为一组(信息码组)，并按一定规则在每一个信息码组之后增加r ($r=n-k$)个校验码元，构成长度为n的新的码字，我们称之为分组码，其中r个监督位的作用是实现检错与纠错。当分组码的信息位和监督位之间为线性关系时，这种分组码就称为线性分组码。线性分组码一般是按照代数规律构造的，故又称为代数编码，一般记为(n，k)码，其中，k为信息码长度，n为码组长度，而$n-k$则为监督码元长度。

线性分组码的主要性质有：①线性分组码具有封闭性，任意两个准用码之和(对于二进制码这个和的含义是模 2 和)仍为一个准用码；②码组间的最小码距等于非零码的最小汉明重量；③全零码必属于线性分组码。

3.3.2 线性分组码的监督关系式和校正子

前面提到的奇偶监督码就是一种最简单的线性分组码，它使用监督码a_0与前面的$a_{n-1}a_{n-2}\cdots a_2a_1$构成监督关系式，也称为监督方程式。接收端为了检测传输过程中是否产生错误，则

$$S=a_{n-1}\oplus a_{n-2}\oplus\cdots\oplus a_0 \tag{3-20}$$

其中S称作校正子，$S=0$表示无错误，$S=1$表示有错误。由于奇偶监督码中只有一位监督码元，只能表示有错和无错。若监督位变为两位，相应可计算两个校正子S_1和S_2，$S_1\ S_2$共有 4 种组合，即 00、01、10、11，可表示 4 种不同信息，除 00 表示无错外，其余 3 种就可表示 3 种不同的错误图样。

推广到一般情况，有r个监督位就可对应r个监督方程式，得到r个校正子，可以用来指示2^r-1个错误图样。若只有 1 位错误，就可以指示2^r-1个错误位置，依此进行错误纠正。对于为(n，k)分组码，要纠正 1 位或 1 位以上的错误，则需要满足$2^r-1>n$。

下面举例说明如何构造线性分组码。

设线性分组码为(n,k)，其中$n=7$，$k=4$，如果要纠正 1 位错误，则要求校验位码元数$r\geqslant 3$，取$r=3$，则$n=k+r=7$，用$a_6a_5a_4a_3a_2a_1a_0$来表示这 7 个码元，用 3 个监督关系式得到 3 个校正子$S_1S_2S_3$，$S_1S_2S_3$与误码位置的关系见表 3-6。

表 3-6 校正子与误码位置

$S_1S_2S_3$	误码位置	$S_1S_2S_3$	误码位置
001	a_0	101	a_4
010	a_1	110	a_5
100	a_2	111	a_6
011	a_3	000	无错



根据表 3-6 所示的对应关系，可以得到如下关系式

$$\begin{cases} S_1 = a_6 + a_5 + a_4 + a_2 \\ S_2 = a_6 + a_5 + a_3 + a_1 \\ S_3 = a_6 + a_4 + a_3 + a_0 \end{cases} \tag{3-21}$$

其中 $a_6a_5a_4a_3$ 为 4 个信息位，它们的取值由输入的信息比特决定，$a_2a_1a_0$ 为监督位，它们的取值根据信息位的取值按监督方程确定。

无错时，$S_1S_2S_3$ 对应 000，即有如下关系式

$$\begin{cases} a_6 + a_5 + a_4 + a_2 = 0 \\ a_6 + a_5 + a_3 + a_1 = 0 \\ a_6 + a_4 + a_3 + a_0 = 0 \end{cases} \tag{3-22}$$

解方程，可得到监督码元计算式

$$\begin{cases} a_2 = a_6 + a_5 + a_4 \\ a_1 = a_6 + a_5 + a_3 \\ a_0 = a_6 + a_4 + a_3 \end{cases} \tag{3-23}$$

因此，给出信息位 $a_6a_5a_4a_3$ 后，根据式(3-23)，就可以计算出监督位 $a_2a_1a_0$，从而得到(7,4)码对应的码组，共有 16 个准用码组，见表 3-7。由上述方程可见，线性分组码中信息位和监督位显然是线性关系。

表 3-7　(7,4)分组码准用码组

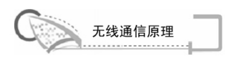
信息位 $a_6a_5a_4a_3$	监督位 $a_2a_1a_0$	信息位 $a_6a_5a_4a_3$	监督位 $a_2a_1a_0$
0000	000	1000	111
0001	011	1001	100
0010	101	1010	010
0011	110	1011	001
0100	110	1100	001
0101	101	1101	010
0110	011	1110	100
0111	000	1111	111

3.3.3　线性分组码的监督矩阵与生成矩阵

1. 监督矩阵

式(3-22)的关系可以写成方程式

$$\begin{cases} 1 \cdot a_6 + 1 \cdot a_5 + 1 \cdot a_4 + 0 \cdot a_3 + 1 \cdot a_2 + 0 \cdot a_1 + 0 \cdot a_0 = 0 \\ 1 \cdot a_6 + 1 \cdot a_5 + 0 \cdot a_4 + 1 \cdot a_3 + 0 \cdot a_2 + 1 \cdot a_1 + 0 \cdot a_0 = 0 \\ 1 \cdot a_6 + 0 \cdot a_5 + 1 \cdot a_4 + 1 \cdot a_3 + 0 \cdot a_2 + 0 \cdot a_1 + 1 \cdot a_0 = 0 \end{cases} \tag{3-24}$$



上式用矩阵形式表示为

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_6 \\ a_5 \\ a_4 \\ a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3-25}$$

或

$$[a_6, a_5, a_4, a_3, a_2, a_1, a_0] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [0,0,0] \tag{3-26}$$

即

$$\boldsymbol{HC}^{\mathrm{T}} = 0^{\mathrm{T}} \text{ 或 } \boldsymbol{CH}^{\mathrm{T}} = 0 \tag{3-27}$$

其中 $\boldsymbol{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$，$\boldsymbol{C} = [a_6, a_5, a_4, a_3, a_2, a_1, a_0]$，$\boldsymbol{C}^{\mathrm{T}}$ 是 $\boldsymbol{C}$ 的转置，$\boldsymbol{H}^{\mathrm{T}}$ 是 $\boldsymbol{H}$ 的转置。这里 $\boldsymbol{H}$ 就称为监督矩阵，$\boldsymbol{C}$ 为码组，信息码与监督码之间的关系完全由监督矩阵 $\boldsymbol{H}$ 确定。可以将 $\boldsymbol{H}$ 分为两部分表示

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & \vdots & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix} = [\boldsymbol{P} \vdots \boldsymbol{I}_r] \tag{3-28}$$

其中 $\boldsymbol{P}$ 为 $r \times k$ 阶矩阵，$\boldsymbol{I}_r$ 为 $r \times r$ 阶方阵，我们将具有 $[\boldsymbol{P} \vdots \boldsymbol{I}_r]$ 形式的矩阵 $\boldsymbol{H}$ 称为典型监督矩阵，根据典型监督矩阵及信息码元很容易计算出各监督码元。

2. 生成矩阵

式(3-23)可写成矩阵形式

$$\begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_6 \\ a_5 \\ a_4 \\ a_3 \end{bmatrix} \tag{3-29}$$



将上式进行转置变换，并将信息码元附在前面，得到

$$[a_6,a_5,a_4,a_3,a_2,a_1,a_0]=[a_6,a_5,a_4,a_3]\begin{bmatrix}1&0&0&0&1&1&1\\0&1&0&0&1&1&0\\0&0&1&0&1&0&1\\0&0&0&1&0&1&1\end{bmatrix} \tag{3-30}$$

即

$$\boldsymbol{C}=\boldsymbol{MG} \tag{3-31}$$

其中 $\boldsymbol{C}=[a_6,a_5,a_4,a_3,a_2,a_1,a_0]$，$\boldsymbol{M}=[a_6,a_5,a_4,a_3]$，$\boldsymbol{G}=\begin{bmatrix}1&0&0&0&1&1&1\\0&1&0&0&1&1&0\\0&0&1&0&1&0&1\\0&0&0&1&0&1&1\end{bmatrix}$，$\boldsymbol{G}$ 称为码的生成矩阵，即由信息码元 $\boldsymbol{M}$ 和生成矩阵 $\boldsymbol{G}$ 可以构造整个码组 $\boldsymbol{C}$ 。

$\boldsymbol{G}$ 矩阵也可以分成两部分

$$\boldsymbol{G}=\begin{bmatrix}1&0&0&0&\vdots&1&1&1\\0&1&0&0&\vdots&1&1&0\\0&0&1&0&\vdots&1&0&1\\0&0&0&1&\vdots&0&1&1\end{bmatrix}=[\boldsymbol{I}_k\vdots\boldsymbol{Q}] \tag{3-32}$$

其中 $\boldsymbol{Q}=\boldsymbol{P}^{\mathrm{T}}$，$\boldsymbol{I}_k$ 为 $k\times k$ 的单位矩阵，具有上式形式的矩阵 $\boldsymbol{G}$ 称为典型生成矩阵，由典型生成矩阵产生的码组一定是系统码，即信息码元不变，而监督码元附在后面。

3.3.4　线性分组码的纠错原理

1. 纠错原理

通信过程中，发送码组会产生误码，根据校正子可以确定错误图样，从而实现纠错。假设发送的码组为 $\boldsymbol{A}$，接收到的码组为 $\boldsymbol{B}$，则收发码组之差为

$$\boldsymbol{E}=\boldsymbol{B}-\boldsymbol{A} \tag{3-33}$$

其中 $\boldsymbol{E}=[e_{n-1},e_{n-2},\cdots,e_0]$，$\boldsymbol{E}$ 中的每一位 e_i 由码组 $\boldsymbol{A}$、$\boldsymbol{B}$ 中的相应位 a_i、b_i 决定。

$$e_i=\begin{cases}0, & a_i=b_i\\1, & a_i\neq b_i\end{cases} \tag{3-34}$$

其中 $e_i=0$ 表示第 i 位无错，$e_i=1$ 表示第 i 位有错。

根据 $\boldsymbol{E}$ 和码组 $\boldsymbol{B}$ 进行译码得到码组 $\boldsymbol{A}$，即

$$\boldsymbol{A}=\boldsymbol{B}-\boldsymbol{E} \tag{3-35}$$

也可写成

$$\boldsymbol{B}=\boldsymbol{A}+\boldsymbol{E} \tag{3-36}$$

无错时，有 $\boldsymbol{AH}^{\mathrm{T}}=0$，有错时，计算得到接收端校正子为

$$\boldsymbol{S}=\boldsymbol{BH}^{\mathrm{T}}=(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{AH}^{\mathrm{T}}+\boldsymbol{EH}^{\mathrm{T}}=\boldsymbol{EH}^{\mathrm{T}} \tag{3-37}$$

可见校正子只与错误图样 $\boldsymbol{E}$ 有关，译码的关键是如何得到错误图样 $\boldsymbol{E}$。因此，如果将接收到的码组 $\boldsymbol{B}$ 按照设计的监督关系进行运算，我们可以得到校正子 $\boldsymbol{S}$，确定错误图样 $\boldsymbol{E}$，

然后将错误图样与接收码组相加，即可纠正错误，并得到发送码组，这就是线性分组码的纠错原理。

2. 监督码元个数与纠错能力的关系

根据线性分组码纠错的基本原理，r 位校正子与错误图样一一对应，对于能纠正 t 个错误的线性分组码应满足不等式

$$2^r = 2^{n-k} \geqslant 1 + C_n^1 + C_n^2 + \cdots + C_n^t = \sum_{i=0}^{t} C_n^i \tag{3-38}$$

其中 C_n^i 表示 n 中取 i 的组合，它表示长度为 n 的码组中有 i 个误码的错误图样数目。

3. 完备码

式(3-38)取等号时表示 r 个校正子与误码不超过 t 个的所有错误图样一一对应，监督位得到了最充分的利用，这种码称作完备码，它表示为

$$2^r = 1 + C_n^1 + C_n^2 + \cdots + C_n^t = \sum_{i=0}^{t} C_n^i \tag{3-39}$$

例如上面介绍的(7,4)码就是完备码，(23,12)也是完备码，(31,21)不是完备码。

4. 汉明码

纠正一位错误的线性分组码称为汉明码，二进制汉明码可以表示为

$$(n,k) = \left(2^m - 1, 2^m - 1 - m\right) \tag{3-40}$$

汉明码的特点是码长 $n = 2^m - 1$，信息位数 $k = 2^m - 1 - m$，监督位数为 $r = m$，最小码距 $d_{\min} = 3$，纠错能力 $t = 1$。如果要产生一个系统的汉明码，可以将矩阵转换成典型监督矩阵，进一步利用 $\boldsymbol{Q} = \boldsymbol{P}^{\mathrm{T}}$ 的关系，得到相应的生成矩阵 $\boldsymbol{G}$。当 r =3、4、5 时，线性分组码(7,4)、(15,11)、(31,26)……都是汉明码。汉明码的编码效率 $R = \dfrac{k}{n} = \dfrac{n-r}{n} = 1 - \dfrac{r}{n}$，当 n 较大时，编码效率接近于 1，所以汉明码是一种高效率的纠错码。

3.4 循 环 码

3.4.1 循环码的基本概念

循环码是一种特殊的线性分组码，它具有许多特殊的代数性质。循环码可用带反馈的移位寄存器实现，实现简单且性能较好，不但可以用于纠正独立的随机错误，也可以用于纠正突发错误，译码算法也较容易，目前计算机纠错编码系统中所使用的线性分组码大部分都为循环码。

循环码具有两个重要的数学特征：封闭性和循环性。封闭性指循环码的码组中任两个码组之和必定为该码组集合中的一个码组，这一特征也是任意分组码都具备的。循环性指循环码中的任一码组循环移位后仍为该码组集合中的一个码组。表 3-8 给出一种(7,3)循环码的全部码组。



表 3-8　(7,3)循环码举例

码组编号	信息位 $a_6a_5a_4$	监督位 $a_3a_2a_1a_0$	码组编号	信息位 $a_6a_5a_4$	监督位 $a_3a_2a_1a_0$
1	000	0000	5	100	1011
2	001	0111	6	101	1100
3	010	1110	7	110	0101
4	011	1001	8	111	0010

为了用代数理论的方法研究循环码的特性，经常将循环码表示成码多项式的形式，即将一个长度为 n 的码组表示为

$$T(x)=a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+\cdots+a_1x+a_0 \tag{3-41}$$

表 3-8 中(7,3)循环码的任意一个码组都可以表示为

$$T(x)=a_6x^6+a_5x^5+a_4x^4+a_3x^3+a_2x^2+a_1x+a_0 \tag{3-42}$$

系数 $a_i=0,1$，其中 $a_i=0$ 的项常略去不写，$a_i=1$ 的项则只写成 x 的项符号而略去系数，这种多项式称为码多项式，在循环码中一般都用码多项式来表示码字。

3.4.2　循环码的生成多项式和生成矩阵

在循环码中，一个 (n,k) 循环码有 2^k 个不同的码组，我们用 $g(x)$ 表示其中前 $(k-1)$ 位皆为零的码组，则 $g(x),xg(x),x^2g(x),\cdots,x^{k-1}g(x)$ 都是码组，且这 k 个码组是线性无关的，可以用来构成循环码的生成矩阵 $\boldsymbol{G}$，而 $g(x)$ 被称为生成多项式。

在包含 k 个信息位的情况下，利用 k 个线性独立的码组可以得到循环码的生成矩阵 $\boldsymbol{G}$

$$\boldsymbol{G}(x)=\begin{bmatrix} x^{k-1}g(x) \\ x^{k-2}g(x) \\ \vdots \\ xg(x) \\ g(x) \end{bmatrix} \tag{3-43}$$

其中 $g(x)=x^r+a_{r-1}x^{r-1}+\cdots+a_1x+1$。因此，一旦生成多项式 $g(x)$ 确定以后，该循环码的生成矩阵就可以确定，进而该循环码的所有码字也就随之确定。例如，表 3-8 给出的循环码中 $n=7$，$k=4$，$n-k=3$，唯一的一个 $(n-k)$ 次多项式代表的码组是第二个码组 0010111，相对应的生成多项式为 $g(x)=x^4+x^2+x+1$，将其代入生成矩阵 $\boldsymbol{G}$ 的表达式中，则得到

$$\boldsymbol{G}(x)=\begin{bmatrix} x^2g(x) \\ xg(x) \\ g(x) \end{bmatrix}=\begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{3-44}$$

将上式变为典型生成矩阵

$$\boldsymbol{G}'(x)=\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{3-45}$$

根据 $\boldsymbol{G}'$ 就可以由 3 个信息位得到如表 3-8 所示的全部码组。

产生循环码的关键是生成矩阵的求解，求生成矩阵的关键又在于寻找一个适当的生成多项式 $g(x)$，例如，对于(7,3)循环码，已知信息位 $a_6a_5a_4$，则可写出循环码组

$$
\begin{aligned}
T(x) &= [a_6, a_5, a_4]\boldsymbol{G}(x) = [a_6, a_5, a_4]\begin{bmatrix} x^2g(x) \\ xg(x) \\ g(x) \end{bmatrix} \\
&= a_6x^2g(x) + a_5xg(x) + a_4g(x) \qquad (3\text{-}46) \\
&= [a_6x^2 + a_5x + a_4]g(x)
\end{aligned}
$$

上式表明，所有码多项式 $T(x)$ 都可以被 $g(x)$ 整除，而且任一幂次不大于 $(k-1)$ 的多项式乘以 $g(x)$ 都是码多项式。

那么如何寻找任意一个(n,k)循环码的生成多项式 $g(x)$？

任一循环码多项式 $T(x)$ 都是 $g(x)$ 的倍式，可写成

$$T(x) = h(x)g(x) \tag{3-47}$$

其中 $h(x) = x^k + h_{k-1}x^{k-1} + \cdots + h_1x + 1$ 是监督多项式，生成多项式 $g(x)$ 本身也是一个码组，而且有

$$T'(x) = g(x) \tag{3-48}$$

码组 $T'(x)$ 是一个 $(n-k)$ 次多项式，所以 $x^kT'(x)$ 是一个 n 次多项式，而且 $x^kT'(x)$ 在模 (x^n+1) 运算下亦为一码组，可写成

$$\frac{x^kT'(x)}{x^n+1} = Q(x) + \frac{T(x)}{x^n+1} \tag{3-49}$$

式(3-49)左边的分子和分母都是 n 次多项式，所以商 $Q(x)=1$，上式简化为

$$x^kT'(x) = (x^n+1) + T(x) \tag{3-50}$$

将式(3-47)和式(3-48)代入式(3-50)，并化简得

$$x^n + 1 = g(x)\left[x^k + h(x)\right] \tag{3-51}$$

式(3-51)表明，长为 n 的循环码的生成多项式 $g(x)$ 必定是 (x^n+1) 中的一个 $(n-k)$ 次因式。

例如，x^7+1 可以分解为：$x^7+1=(x+1)(x^3+x^2+1)(x^3+x+1)$，则(7,3)循环码的生成多项式 $g(x)$ 必定是上式中一个 $(n-k)=4$ 次的因子，且可以看出这样的因子有两个，分别为

$$(x+1)(x^3+x^2+1) = x^4+x^2+x+1 \tag{3-52}$$

$$(x+1)(x^3+x+1) = x^4+x^3+x^2+1 \tag{3-53}$$

以上两式都可以作为生成多项式，显然，选用不同的多项式得到的循环码组也不同。

3.4.3 循环码的编码原理

循环码的编码是根据给定的(n,k)值选择生成多项式 $g(x)$，即从 (x^n+1) 的因子中选择一个 $(n-k)$ 次多项式作为 $g(x)$，$g(x)$ 的最高次幂为 $r=n-k$，这恰好是码组中校验码元数。根据所有码多项式 $T(x)$ 都可以被 $g(x)$ 整除的原则，就可以对给定的信息位进行编码。



如果输入信息序列用多项式 $I(x)$ 表示，由于信息码元为 k 个，所以 $I(x)$ 的最高次幂为 $k-1$，进行如下除法运算

$$\frac{x^{n-k}I(x)}{g(x)}=Q(x)+\frac{R(x)}{g(x)} \tag{3-54}$$

式中 $x^{n-k}I(x)$ 的最高次幂为 $(n-k)+(k-1)=n-1$，商式 $Q(x)$ 的最高次幂为 $(n-1)-(n-k)=k-1$，余式 $R(x)$ 的最高次幂为 $n-k-1=r-1$，即 $Q(x)$ 的幂次与码组中监督码元数减 1 相对应，$R(x)$ 的幂次与码组中校验码元数减 1 相对应。因此，式(3-54)又可以写为

$$x^{n-k}I(x)=Q(x)g(x)+R(x) \tag{3-55}$$

移项后得到

$$C(x)=x^{n-k}I(x)+R(x)=Q(x)g(x) \tag{3-56}$$

$C(x)$ 就是经过除法运算后所编成的循环码的多项式表示。

综上，循环码编码步骤归纳如下。

(1) 用 x^{n-k} 乘 $I(x)$，在信息码后面加上 $r=n-k$ 个“0”。

(2) 用生成多项式 $g(x)$ 除 $x^{n-k}I(x)$，得到商式和余式，即 $\dfrac{x^{n-k}I(x)}{g(x)}=Q(x)+\dfrac{R(x)}{g(x)}$。

(3) 编出码组，即联合 $R(x)$ 和 $x^{n-k}I(x)$ 得到码多项式 $C(x)=x^{n-k}I(x)+R(x)$。

【例 3.4】 对信息序列 $I(x)=11101010001=x^{10}+x^9+x^8+x^6+x^4+1$ 进行(15,11)循环编码，生成多项式 $g(x)=x^4+x+1$，$n=15$，$k=11$。

解： $x^{n-k}I(x)=x^{15-11}(x^{10}+x^9+x^8+x^6+x^4+1)=x^{14}+x^{13}+x^{12}+x^{10}+x^8+x^4$

利用长除法运算得到

$$\frac{x^{n-k}I(x)}{g(x)}=Q(x)+\frac{R(x)}{g(x)}=(x^{10}+x^9+x^8+x^7+x^6+x^4+x^2+x)+\frac{x^3+x}{x^4+x+1}$$

则商式和余式分别为

$$Q(x)=x^{10}+x^9+x^8+x^7+x^6+x^4+x^2+x$$

$$R(x)=x^3+x$$

故所编成的码组为

$$C(x)=x^{n-k}I(x)+R(x)=x^{14}+x^{13}+x^{12}+x^{10}+x^8+x^4+x^3+x$$

系统输出的循环码组写成一般数字序列为 111010100011010。根据循环码的定义式，可以用多项式的除法运算器来实现编码，实际上是在给定 $g(x)$ 下，对输入信息序列进行除法运算。

在接收端对循环码进行译码时，通常有两个要求：检错和纠错。达到检错目的的译码相对简单，主要通过判断接收到的码组 $P(x)$ 是否能被生成多项式 $g(x)$ 整除，当传输中没有发生错码时，接收码组能被 $g(x)$ 整除。若有差错，接收码组必不能被 $g(x)$ 整除。因此，根据余式是否为零就可以判断码字中有无错误。需要注意的是接收码组存在错码时也有可能被整除，这时错码将不能被检出。

循环码的译码可以简单概括为以下 3 步。

(1) 由接收到的码多项式 $P(x)$ 计算校正子(伴随式)多项式 $R_E(x)$。

(2) 由校正子 $R_E(x)$ 确定错误图样 $E(x)$。

(3) 将错误图样 $E(x)$ 与 $P(x)$ 相加，纠正错误。

第 1 步运算和检错译码类似，也就是求解 $P(x)$ 整除 $g(x)$ 的余式，第 3 步也很简单。因此，纠错码译码器的复杂性主要取决于译码过程的第 2 步。

3.4.4 几种重要的循环码

1. BCH 码

BCH 码由 R. C. Bose、D. K. Chaudhuri 和 A. Hocquenghem 共同提出，并因此而得名。BCH 码是一种能纠正多位错误的循环码，它可以根据纠错能力的要求来直接确定码的构造，它的参量可以在大范围内变化，选用灵活，适用性强，是一类应用广泛的差错控制编码。

BCH 码分为本原 BCH 码和非本原 BCH 码两类。最常用的 BCH 码是本原 BCH 码，对于任意正整数 m ($m \geqslant 3$)和 t ($t < (2^m - 1)/2$)，本原 BCH 码的主要参量及关系式为：分组码长 $n = 2^m - 1$，信息码位数 $k \geqslant n - mt$，最小汉明距离 $d_{\min} \geqslant 2t + 1$。

BCH 码可以纠正或检出 t 位错误，并可以提供灵活的参量选择，码长可达到上百比特，因此它是目前同样码长和码率的所有分组码中的最优码。然而，求 BCH 码的生成多项式是一项非常烦琐的工作。前人已将 BCH 码生成多项式的研究结果列成表(见表 3-9 和表 3-10)，通过查表可以得到不同 n、k、t 取值情况下 BCH 码的生成多项式 $g(x)$，工程设计中直接查表即可使用。

表 3-9 中生成多项式的系数都是用八进制表示的，例如，$n=15$，$k=11$，$t=1$ 的 BCH 码生成多项式为 $g(x)=(23)_8$，将八进制数表示成二进制数得到 $g(x)=(23)_8=(010011)_2$，它对应了生成多项式各项的系数，即 $g(x)=x^4+x+1$。表 3-10 中，(23,12)是一个非本原 BCH 码，它的码距为 7，可以纠正 3 个随机独立的错误，其生成多项式为 $g(x)=(5343)_8=(101011100011)_2=x^{11}+x^9+x^7+x^6+x^5+x+1$。在实际应用中，BCH 码的码长都为奇数，有时为了得到偶数码长，可将 BCH 码的生成多项式乘以一个因子 $x+1$，它相当于在原 BCH 码上增加了一个校验位，从而得到更强的纠错能力。

表 3-9　部分本原 BCH 码的参数

n	k	t	生成多项式 $g(x)$的系数(八进制)
15	7	2	721
	5	3	2461
31	21	2	3551
	16	3	107657
	11	5	5423325
63	51	2	12471
	45	3	1701317
	39	4	166623567
	30	6	157464165547

续表

n	k	t	生成多项式 $g(x)$的系数(八进制)
127	113	2	41567
	106	3	11554743
255	239	2	267543
	231	3	156720665
511	493	2	1132353

表 3-10　部分非本原 BCH 码的参数

n	k	t	m	生成多项式 $g(x)$的系数(八进制)
17	9	2	8	727
21	12	2	6	1663
23	12	3	1	5343
33	22	2	10	5145
33	12	4	10	3777
41	21	4	20	6647133
47	24	5	23	43073357
65	3	2	12	10761
65	40	4	12	354303067
73	46	4	9	1717773537

2. CRC 码

循环冗余校验码(CRC 码)是一种非常适于检错的差错控制码。它有两个突出优点：①检错能力强，可以检出多种可能的组合性错误，对随机错误和突发错误都能以较低冗余度进行严格检验；②比较容易实现编、解码电路，在计算机通信中应用广泛。

CRC 码可以检测出以下几种形式的错误。

(1) 突发长度不超过 $n-k$ 的全部错误。

(2) 大部分突发长度等于 $n-k+1$ 的错误，其中不可检的这类错误为 $2^{-(n-k-1)}$ 。

(3) 大部分突发长度大于 $n-k+1$ 的错误，其中不可检的这类错误为 $2^{-(n-k)}$ 。

(4) 所有与需用码组距离小于等于 $d_{\min}-1$ 的错误。

(5) 所有奇数个随机错误。

表 3-11 给出了已成为国际标准的常用的 4 种 CRC 码，其中 CRC-12 用于长度为 6 比特的字符，另外 3 种码则常用于 8 比特字符。

表 3-11 国际标准的 4 种 CRC 码

CRC 码	生成多项式 $g(x)$
CRC-12 码	$x^{12}+x^{11}+x^3+x^2+x+1$
CRC-16 码	$x^{16}+x^{15}+x^2+1$
CRC-CCITT 码	$x^{16}+x^{12}+x^5+1$
CRC-32 码	$x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^8+x^7+x^5+x^4+x^2+x+1$

3. RS 码

RS 码是 Reed-Solomon 的缩写，又称里所码，它是一类非二进制 BCH 码，纠错能力很强。在(n,k)RS 码中，输入信号每 kn 比特分成一组，每组包括 k 个符号，每个符号由 m 比特组成。一般 RS 码常用 $m=8$ 比特，8 比特 RS 码具有很大应用价值，可以纠 t 个符号错误的 RS 码参量如下。

(1) 码长：$n=2^m-1$ 符号或 $m(2^m-1)$ 比特。

(2) 信息段：k 符号或 km 比特。

(3) 监督段：$n-k=2t$ 符号或 $m(n-k)$ 比特。

(4) 最小汉明距离：$d_{\min}=2t+1$ 符号或 $m(2t+1)$ 比特。

例如，(255,223)RS 码表示码块长度共有 255 个符号，其中信息段有 223 个符号，监督段有 32 个检验符号。在这个由 255 个符号组成的码块中，可以纠正在这个码块中出现的 16 个分散的或者 16 个连续的错误符号。RS 码是一种纠错能力很强的纠错码，特别适用于存在突发错误的信道如移动通信衰落信道和多进制调制的场合，目前 RS 码被广泛应用于各种商业用途，如数字通信和数字存储领域及卫星通信中，它也被用于 DSL 和 WiMAX，广播系统中 DVB 和 ATSC 也闪现着它的身影，在电脑科学里，它是第六层标准 RAID 的重要成员。

3.5 卷 积 码

3.5.1 卷积码的基本概念

分组码是把 k 个信息比特的序列编成 n 个比特的码组，每个码组的 $(n-k)$ 个校验位仅与本码组的 k 个信息位有关，而与其他码组无关。为了达到一定的纠错能力和编码效率，分组码的码组长度一般都比较大，编译码时必须把整个信息码组存储起来，而由此产生的译码时延随 n 的增大而增加。与分组码不同，卷积码的 n 位编码，不仅与当前段的 k 个信息位有关，还与前面 m 段的信息位有关，整个编码过程可以看成是输入信息序列与由移位寄存器和模 2 和连接方式所决定的另一个序列的卷积，卷积码也由此得名。

卷积码常表示为(n,k,m)，n 表示卷积码编码器输出端码元数，k 表示编码器输入端信息位，m 表示编码器中寄存器的节数，卷积码的 k 和 n 通常很小。从编码器的输入端来看，卷积码仍以 k 位数据为一组，分组输入。从输出端看，卷积码是非分组的，它输出的 n

位码元不仅与当前输入的 k 个信息位有关，而且还与前面 m 段的输入信息位有关，所以卷积码属于有记忆的非分组码。卷积码为有记忆编码，其记忆或称约束度为 $N=m+1$，编码过程中互相关联的码元个数为 nN，称 nN 为编码的约束长度。卷积码的纠错能力随 N 的增大而提高，而差错率随 N 的增大呈指数下降。在编码器复杂性相同的情况下，卷积码的性能优于分组码，但卷积码没有分组码那样的严密数学分析，需要通过计算机进行优码的搜索。卷积码的码长较小，因此适合于串行形式传输，且延时小。

3.5.2　卷积码的编译码原理

1. 卷积码编码器

卷积码编码器的一般结构形式如图 3.11 所示，它包括几大部分：① N 段输入移位寄存器，每段移位寄存器有 k 级存储单元，共 Nk 位寄存器；②一组 n 个模 2 加法器，n 等于卷积编码输出位数，每个模 2 加法器连接到一些移位寄存器的输出端，数目可以不同，连接的移位寄存器也可以不同；③ n 位输出移位寄存器，n 个模 2 加法器与 n 位输出移位寄存器一一对应连接，模 2 加法器的运算结果即为卷积编码输出，每输入 k 个比特，得到 n 个输出比特。

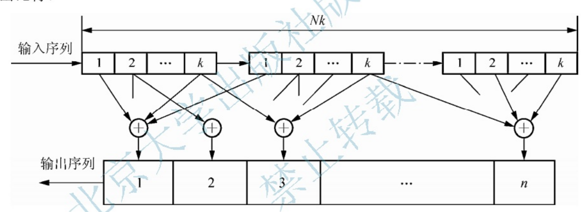


图 3.11　卷积码编码器的一般结构

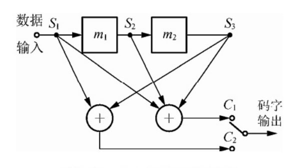


图 3.12　(2,1,2)卷积码编码器

图 3.12 所示为一个(2,1,2)卷积码编码器，图中 $n=2$，$k=1$，$m=2$，它由移位寄存器、模 2 加法器和开关电路组成。每输入一个信息比特经编码器产生两个输出比特。假设移位寄存器的起始状态为全零，即 $S_1S_2S_3$ 为 000，S_1 为当前输入数据，而移位寄存器状态 S_2S_3 存储之前的数据，输出码字由下式确定

$$\begin{cases} C_1 = S_1 \oplus S_2 \oplus S_3 \\ C_2 = S_1 \oplus S_3 \end{cases} \tag{3-57}$$

起始状态，各移位寄存器为全零，当第一个输入比特为 0 时，输出比特为 00；若输入比特为 1，则输出比特为 11；当第二个比特进入时，第一个比特右移一位，此时输出比特受到当前输入比特和前面输入比特的影响；第三个比特到来时，第一、二个输入比特分别右移一位，此时输出比特受到当前输入比特和前面两个输入比特的影响；第四个比特到来时，第一个输入比特移出寄存器而消失，此时输出比特受到当前输入比特和前面两个输入比特的影响。

表 3-12 给出了(2,1,2)卷积码编码器在移位寄存器的起始状态为全零、编码器输入序列为 11010000 时，对应的编码器输出移位寄存器的状态变化过程，其中输入移位寄存器 S_2S_3 的 4 种状态 00、01、10、11 分别用 a、b、c、d 表示。(2,1,2)卷积码的编码约束度为 3，约束长度为 6。

表 3-12　(2,1,2)卷积码编码器的状态变化过程

S_1	1	1	0	1	0	0	0	0
S_2S_3	00	01	11	10	01	10	00	00
C_1C_2	11	01	01	00	10	11	00	00
状态	a	b	d	c	b	c	a	a

2. 卷积码的描述

卷积码的描述方法有两类：图解法和解析法。其中图解法包括：树状图、状态图、网格图；解析法包括：矩阵形式、生成多项式形式。

1) 树状图

上面的(2,1,2)卷积码编码器中随着输入序列的进入，编码器移位的过程和各种可能的输出序列可用图 3.13 所示的树状图来表示。图 3.13 所示的树状图从节点 a 开始，由 a 出发有两条路径，$a=0$ 取上支路，$a=1$ 取下支路，输入第二个比特时，移位寄存器右移一位，上支路的状态仍为 00(a)，下支路状态变为 01(b)。新的比特到来时，随着移位寄存器状态和输入比特的不同，树状图分为 4 个支路，以此类推，可得整个树状图。输入不同的信息序列，编码器就走不同的路径，从而得到输出不同的码序列。例如输入数据为 1101 时，其路径如图中虚线所示，输出码序列为 11010100，与表 3-12 结果相同。

2) 状态图

将已到达稳定状态的一节网格取出，当前状态与下一状态合并，得到图 3.14 所示的状态图。图 3.14 中，有 a、b、c、d 共 4 个节点，同样分别表示 S_2S_3 的 4 种可能状态：00、01、10、11。每个节点有两条线离开该节点，实线表示输入数据为 0，虚线表示输入数据为 1，两个闭合圆圈分别表“$a \to a$”和“$d \to d$”的状态转移，线旁的数字即为输出码字。



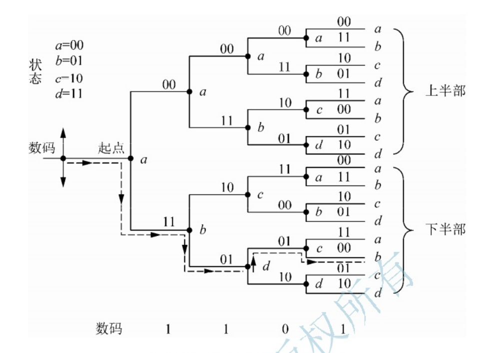


图 3.13　(2,1,2)卷积码的树状图

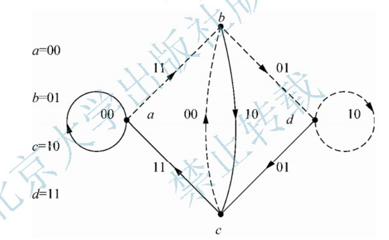


图 3.14　(2,1,2)卷积码的状态图

3) 网格图

将图 3.13 所示的树状图中相同状态的节点合并，可以得到一种简化的图，即网格图，如图 3.15 所示。图中，上支路用实线表示，下支路用虚线表示，每个支路上标注的元素仍为输出比特，节点上标注的为状态。

4) 卷积码的矩阵描述

仍以(2,1,2)卷积码编码器为例，假设移位寄存器起始状态全为 0，若输入信息序列和输出序列分别表示为：$\boldsymbol{M}=[m_1,m_2,m_3,\cdots]$，$\boldsymbol{C}=[c_{1,1},c_{2,1},c_{1,2},c_{2,2},c_{1,3},c_{2,3},\cdots]$。

当第一个信息比特输入时，对应两个输出比特为：$c_{1,1}=m_1$，$c_{1,2}=m_1$。

当第二个信息比特输入时，对应两个输出比特为：$c_{1,2}=m_2+m_1$，$c_{2,2}=m_2$。

当第三个信息比特输入时，对应两个输出比特为：$c_{1,3}=m_3+m_2+m_1$，$c_{2,3}=m_3+m_1$。

当第 j 个信息比特输入时，对应两个输出比特为：$c_{1,j}=m_j+m_{j-1}+m_{j-2}$，$c_{2,j}=m_j+m_{j-2}$，写成矩阵形式为

$$\left[m_j,m_{j-1},m_{j-2}\right]\begin{bmatrix}1&1\\1&0\\1&1\end{bmatrix}=\left[c_{1,j},c_{2,j}\right] \tag{3-58}$$

图 3.15　(2,1,2)卷积码的网格图

卷积码可以由生成矩阵完全描述，生成矩阵与生成多项式之间存在确定的关系，根据前面的(2,1,2)编码器可知其生成序列为

$$g_1=(1,1,1)=\left(g_1^1,g_1^2,g_1^3\right) \tag{3-59}$$

$$g_2=(1,0,1)=\left(g_2^1,g_2^2,g_2^3\right) \tag{3-60}$$

将 g_1、g_2 交错排列，可得生成矩阵

$$\begin{aligned}\boldsymbol{G}&=\begin{bmatrix}g_1^1&g_2^1&g_1^2&g_2^2&g_1^3&g_2^3&&&&&&\\&&g_1^1&g_2^1&g_1^2&g_2^2&g_1^3&g_2^3&&&&\\&&&&g_1^1&g_2^1&g_1^2&g_2^2&g_1^3&g_2^3&&\\&&&&&&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\end{bmatrix}\\&=\begin{bmatrix}G_1&G_2&G_3&&\\&G_1&G_2&G_3&\\&&G_1&G_2&G_3\\&&&\cdots&\cdots&\cdots\end{bmatrix}\end{aligned} \tag{3-61}$$

给定输入信息序列 $\boldsymbol{M}$，根据生成矩阵 $\boldsymbol{G}$，即可得到输出序列 $\boldsymbol{C}=\boldsymbol{MG}$。与上述(2,1,2)卷积码类似，可推广到一般情况，对于(n,k,m)码，有

$$\boldsymbol{M}=\left(m_{1,1},m_{2,1},\cdots,m_{k,1},m_{1,2},m_{2,2},\cdots,m_{k,2},\cdots\right),\quad \boldsymbol{C}=\left(c_{1,1},c_{2,1},\cdots,c_{n,1},c_{1,2},c_{2,2},\cdots,c_{n,2},\cdots\right)$$

已知该码的生成序列多项式为

$$g_{i,j}=\left(g_{i,j}^1,g_{i,j}^2,g_{i,j}^3,\cdots,g_{i,j}^l,\cdots,g_{i,j}^N\right),\ i=1,2,\cdots,k;\ j=1,2,\cdots,N \tag{3-62}$$



其中 $g_{i,j}$ 表示了每组 k 个输入比特中第 i 个比特经 $l-1$ 组延迟后的输出与每组 k 个输出比特中第 j 个模 2 和的输入端的连接关系，$g_{i,j}=1$ 表示有连线，$g_{i,j}=0$ 表示无连线，因此生成矩阵的一般形式为

$$\boldsymbol{G}=\begin{bmatrix} G_1 & G_2 & G_3 & \dots & G_N & & & \\ & G_1 & G_2 & G_3 & \dots & G_N & & \\ & & G_1 & G_2 & G_3 & \dots & G_N & \\ & & & \cdots & \cdots & \cdots\cdots & \cdots \end{bmatrix} \tag{3-63}$$

其中 $\boldsymbol{G}_l\ (l=1,2,\cdots,N)$ 是 k 行 n 列子矩阵

$$\boldsymbol{G}_l=\begin{bmatrix} g_{1,1}^l & g_{1,2}^l & g_{1,3}^l & \cdots & g_{1,n}^l \\ g_{2,1}^l & g_{2,2}^l & g_{2,3}^l & \cdots & g_{2,n}^l \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{k,1}^l & g_{k,2}^l & g_{k,3}^l & \cdots & g_{k,n}^l \end{bmatrix} \tag{3-64}$$

3. *卷积码译码*

卷积码译码可分为代数译码和概率译码两大类。代数译码是利用生成矩阵和监督矩阵来译码，最主要的方法是大数逻辑译码，该方法利用编码本身的大数结构，而不考虑信道的统计，硬件实现简单，但性能较差。概率译码比较实用的有维特比译码和序列译码两种，其中维特比译码应用较多，这里只简单介绍维特比译码方法。

维特比算法是 1967 年由 Viterbi 提出的一种卷积码的译码算法，该算法的本质是利用最大似然准则进行译码，其基本思想是将接收码字序列与所有可能的发送码字序列进行比较，从中选择与接收序列汉明距离最小的发送序列(汉明距离最小就相当于似然函数最大)作为译码输出。对于长度为 L 的二进制发送序列，需要对可能发送的 2^L 个不同的序列的 2^L 条路径似然函数累加值进行比较，选取其中最大的一条。显然，译码算法的计算量将随着 L 的增加呈指数增加，实际中需要采取一定措施简化。维特比译码使用网格图描述卷积码，每个可能的发送序列都与网格图中的一条路径相对应。根据网格图的路径汇聚特性，如果在某个节点上发现某条路径已不可能与接收序列的距离最小，那么就放弃这条路径，然后在剩下的“幸存”路径中重新选择码路径，这样一直进行到最后第 L 级。由于这种方法过早丢弃了那些不可能的路径，因而减轻了译码的工作量。

维特比译码算法的基本步骤：对于网格图第 i 级的每个节点，计算到该节点的所有路径的距离量度，即在前面 $i-1$ 级路径距离量度的基础上累加上第 i 条支路的距离量度，并从中选择距离量度最小的路径作为幸存路径，其他路径丢弃。上述译码过程可概括为“累加—比较—选择(ACS)”运算。

为了具体阐述维特比译码的过程，这里仍以(2,1,2)卷积码为例进行分析。设编码器输入信息为 11010000，由编码器输出的码字为 1101010010110000，接收端接收到的码序列为 0101011010010010，有 4 位码元发生错误，图 3.16 的格图说明了整个译码过程。

图 3.16 中先选前 3 个码作为标准，对到达第 3 级的 4 个节点的 8 条路径进行比较，逐步算出每条路径与接收码字之间的累加码距。累加码距分别用括号内数字标出，对照后，

保留一条到达该节点的码距最小的路径作为幸存路径。再将当前节点移到第 4 级，计算、比较、保留幸存路径，直至最后得到到达终点的一条幸存路径，即为译码路径，并根据该路径得到译码结果，如图 3.16 中实线所示。

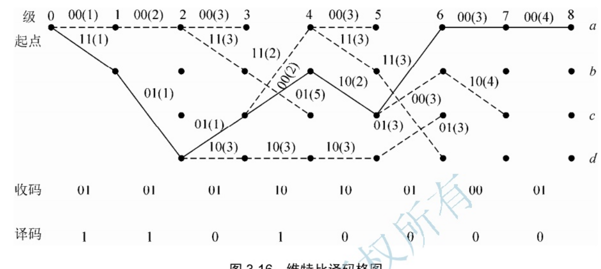


图 3.16　维特比译码格图

维特比译码需要在网格图的每一列节点处进行累计距离的“累加—比较—选择”运算，并带有译码的回溯过程。译码器应由软、硬件联合组成，译码器的复杂性也将随着状态数和约束长度的增加而上升。根据上面的讨论，可以得出维特比译码器的基本结构如图 3.17 所示。图 3.17 中，网格图参量存储器用于寄存可能的发码路径参数，输入变换和支路量度计算将输入信号变换为可用计算距离的量度，并计算收码与可能的发码路径间距离，然后由 ACS 电路完成支路量度的“累加—比较—选择”作用，将计算结果存入路径量度存储器，再进行最佳路径选择，利用输出缓冲器完成回溯译码过程。

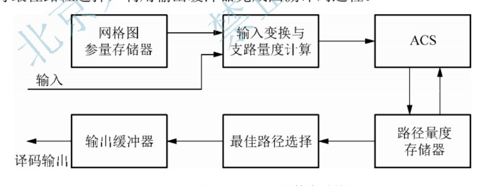


图 3.17　维特比译码器的基本结构

3.6　Turbo 码

3.6.1　Turbo 码的基本概念

Turbo 码又称并行级联卷积码，它是由法国的 C. Berrou 等于 1993 年在国际会议 ICC 上提出的一种新型信道纠错编码方案。Turbo 码巧妙地将卷积码和随机交织器结合在一起，能够产生很长的码字并提供更好的传输性能，达到接近随机编码的目的。Turbo 码一经提

出，就以其优异的性能和相对简单可行的编译码算法吸引了众多研究者的目光。

香农极限定理指出，要想在一个存在噪声的确定带宽的信道中可靠地传送信号，可以加大信噪比或在信号编码中加入附加的纠错码；要想提高信号编码效率达到信道容量，就要使编码的长度尽可能长而且尽可能随机，但这在实际中因为计算量太大而不可能实现。Turbo 码的不同之处在于它通过一个交织器，使之达到接近香农极限的性能。此外，它采用迭代译码策略来逼近最大似然译码，使得译码复杂性大大降低。

3.6.2　Turbo 码的编译码原理

1. *Turbo 码的编码原理*

Turbo 码编码器由两个或两个以上的简单分量编码器(也称子编码器)通过交织器并行级联在一起构成，结构如图 3.18 所示，图中两个分量编码器分别采用递归系统卷积码(RSC)编码器，且具有完全相同的结构。交织器是一个单输入单输出设备，由一定数量的存储单元构成，$M \times N$ 存储单元构成存储矩阵，其中 M 为存储矩阵的行数，N 为存储矩阵的列数，各个存储单元可用它在矩阵中所处的行数和列数来表示。交织器的输入与输出符号序列有相同的字符集，只是各符号在输入与输出序列中的排列顺序不同。编码器将输入数据的 n 个比特分为一组，由编码器 1 进行行编码，再经过交织器由编码器 2 进行列编码。对两路编码的校验位进行抽取，删除适当码元，以提高码率。对于卷积码由于其码率较低，可进行抽取，对于分组码由于其码率较高，可直接省略。最后进行复接，完成并串转换。

交织器的引入可以说是 Turbo 码的一大特色，信息比特流顺序流入交织器，以某种方式乱序读出，或者以乱序的形式读入，再以顺序的形式读出，这种决定读出顺序的方法称为交织器的交织方法。交织器的作用一般是对抗突发错误，但更重要的是它可以改变码的重量分布，将原始信息序列置乱，使交织前后的信息序列的相关性减弱。交织长度越长，相邻反馈信号的相关性就越低，可更好地实现迭代译码。

交织器的设计准则：最大程度地置乱原数据排列顺序，避免置换前相距较近的数据在置换后仍相距较近，特别要避免置换前相邻数据在置换后再次相邻；尽可能避免与同一信息位直接相关的两个分量编码器中的校验位在复用时均被删除；对于不归零的编码器，交织器设计时要避免出现错误平层效应。在满足上述要求的交织器中再选择一个好的交织器，使码字之间的最小距离 $d_{\min}$ 尽可能大，而重量为 $d_{\min}$ 的码字数要尽可能少，以改善 Turbo 码在高信噪比时的性能。

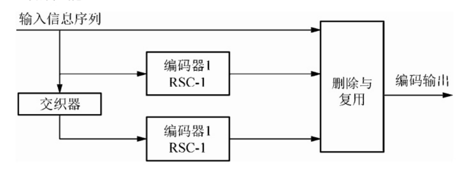


图 3.18　Turbo 码编码器结构

2. Turbo 码的译码

Turbo 码的译码过程采用迭代反馈的方法，每次迭代采用的是软输入和软输出，并不断地将输出的码元反馈到输入，与输入进行比较以增加译码的正确性。Turbo 码译码器结构如图 3.19 所示，由两个软输入软输出(SISO)译码器 DECl 和 DEC2 串行级联组成，交织器与编码器中所使用的交织器相同。Turbo 码的译码过程：译码器 DECl 对分量编码器 RSCl 输出的校验码进行最佳译码，产生关于信息序列中每一比特的似然信息，并将其中的“新信息”经过交织送给 DEC2，译码器 DEC2 将此信息作为先验信息，对分量编码器 RSC2 输出的校验码进行最佳译码，产生关于交织后的信息序列中每一比特的似然比信息，然后将其中的“新信息”经过解交织送给 DECl，进行下一次译码。这样，经过多次迭代，DECl 或 DEC2 的新信息趋于稳定，似然比渐近值逼近于对整个码的最大似然译码，然后对此似然比进行判决，即可得到信息序列的每一比特的最佳估值序列。

Turbo 码译码器采用迭代译码，最佳译码策略是根据接收序列计算后验概率，为降低长码计算的复杂度，由两个分量码译码器分别计算后验概率。常见的 Turbo 码的译码算法有 MAP 算法(最大后验概率)、BCJR 算法、对数域的 LOG-MAP 算法及 MAX-LOG-MAP 算法、减少状态搜索的 M-BCJR 和 T-BCJR 算法、滑动窗 BCJR 算法和只有前向递归的 OSA 算法、SOVA 算法及其改进算法等。

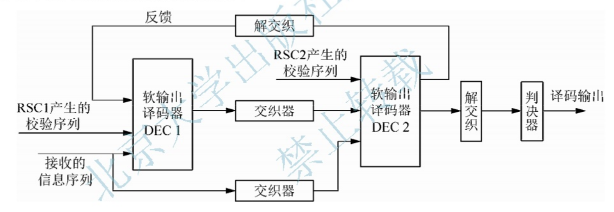


图 3.19 Turbo 码译码器的结构

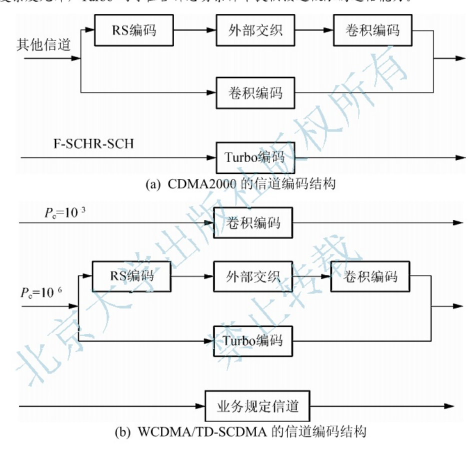
Turbo 码的优点是编码性能高，编码后的速率接近香农容量极限，同时编码的可靠性增强，但其缺点是时延较大，因此 Turbo 编码适用于对误块率要求严格，但时延要求不严格的场合，如 3G 中高速下载类业务。在 3G 移动通信、WiMAX、深空通信等许多国际标准中，Turbo 码都被作为首推的纠错编码。在第 2 代移动通信系统中应用卷积编码和交织，对保证话音和低速数据业务的业务质量有很好的效果。第 3 代系统比第 2 代的业务种类大大增加，对信道编码提出了更高的要求，设计信道编码方案，不仅要从用户业务的要求考虑，如信息的准确度、允许的时延等，也应从提高系统增益的全局优化的角度，与分集接收、改进调制解调方法、系统的经济性等其他因素综合考虑。从目前的研究来看，Turbo 码在学术界研究中的地位相当重要，Turbo 码与空时码、TCM 的结合、Turbo 码均衡技术、Turbo 码多用户检测技术以及在 MIMO 信道、协作通信中的应用均为当前的研究热点。



应用案例

Turbo 码在移动通信中的应用

Turbo 码以其强大的纠错能力和灵活性，在第 3 代移动通信系统中得到广泛的应用。目前，WCDMA、CDMA2000 以及 TD-SCDMA 等标准的第三代移动通信系统都选择将 Turbo 码作为信道编码方案的一个重要组成部分(相应的信道编码结构如图 3.20 所示)。由于 Turbo 码的译码采用了迭代译码的思想，因此可以将它推广到 CDMA 多用户检测中，实现基于 Turbo 码译码原理的 CDMA 多用户检测接收。可以说，只要时延和复杂度允许，Turbo 码可在各种恶劣条件下提供接近极限的通信能力。



(a) CDMA2000 的信道编码结构

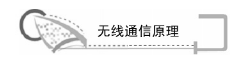
(b) WCDMA/TD-SCDMA 的信道编码结构

图 3.20　Turbo 码在移动通信中的应用

3.7　LDPC 码

3.7.1　LDPC 码的基本概念

LDPC 码的全称是低密度奇偶校验码(Low Density Parity Check Codes，LDPC)，它是一种线性分组码。LDPC 码最初是由 Gallager 在 1962 年提出的，但在 LDPC 码提出之初并未受到人们的重视。后来直到 1996 年 Mackay 等人发现 LDPC 码具有更低的线性译码复杂度，其优秀性能也可以逼近香农极限，LDPC 码才得到普遍关注和研究应用。LDPC 码与 Turbo 码同属于复合编码，两者性能接近，但 LDPC 码比 Turbo 码译码更简单，更易实现，



而且基于规则双向图的 LDPC 码的性能可以优于 Turbo 码。

Gallager 对 LDPC 码的定义为：LDPC 码是一类线性分组码，设码长为 n，信息位为 k，监督位为 $r=n-k$，则它的奇偶校验矩阵 $\boldsymbol{H}$ 是一个 $r\times n$ 的矩阵。

二进制LDPC码的校验矩阵满足每一行和列中1的个数是固定的，其中每一列 $j\,(j>3)$ 个 1，每一行 $i\,(i>j)$ 个 1，列之间 1 的重叠数目小于等于 1，并且校验矩阵 $\boldsymbol{H}$ 中大部分元素为零，元素 1 的密度非常低，所以 LDPC 码的校验矩阵 $\boldsymbol{H}$ 实际上是一个几乎全部由 0 组成的稀疏矩阵，该码也由此得名。LDPC 码一般也表示为 (n,j,i)，下式给出了一个简单的 LDPC 码(10,2,4)的校验矩阵 $\boldsymbol{H}$。

$$\boldsymbol{H}=\begin{bmatrix}1&1&1&1&0&0&0&0&0&0\\1&0&0&0&1&1&1&0&0&0\\0&1&0&0&1&0&0&1&1&0\\0&0&1&0&0&1&0&1&0&1\\0&0&0&1&0&0&1&0&1&1\end{bmatrix} \tag{3-65}$$

式(3-65)中的校验矩阵是一个系数矩阵，可见，它对应的 LDPC 码的码长为 $n=10$，每列 2 个 1($j=2$)，每行 4 个 1($i=4$)。

LDPC 码的校验矩阵也可以利用图论中的二分图或双向图表示，图的下方有 n 个节点，每个节点表示码字的信息位，对应于校验矩阵的各列，图的上方有 r 个节点，每个节点表示码字的一个校验集，称为校验节点，代表校验方程，对应于校验矩阵各行，与校验矩阵中“1”元素相对应的左右两节点之间存在连接边，这两条边两端的节点称为相邻节点，每个节点相连的边数称为该节点的度数。由于该方法是由 Tanner 最先提出，因此人们又将这种图模型称之为 Tanner 图，图 3.21 给出了(10,2,4)LDPC 码的 Tanner 图。

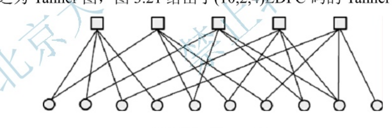


图 3.21　Tanner 图

3.7.2　LDPC 码的编译码原理

LDPC 码编码时先通过矩阵变换得到一个生成矩阵，可以采用高斯消元法将校验矩阵 $\boldsymbol{H}$ 变换为系统形式的校验矩阵 $\boldsymbol{H}'=[\boldsymbol{P}:\boldsymbol{I}]$，由于矩阵 $\boldsymbol{H}$ 的结构是随机的，很可能在变换得到 $\boldsymbol{H}'$ 的过程中需要对矩阵 $\boldsymbol{H}$ 进行列交换操作，并记录列交换信息。根据 $\boldsymbol{H}'$，得到其对应的生成矩阵 $\boldsymbol{G}'=\begin{bmatrix}\boldsymbol{I}\\\boldsymbol{P}\end{bmatrix}$。然后利用信息序列与生成矩阵相乘实现编码，若输入信息序列为 $\boldsymbol{X}$，则编码后的输出序列为 $\boldsymbol{C}'=\boldsymbol{G}'\boldsymbol{X}$。

LDPC 码具有良好性能的重要原因是 LDPC 码采用了基于置信传播的迭代译码算法，这是一种迭代概率译码方法，采用逐符号、软输入软输出(SISO)的译码。对于每个符号的可靠性的量度可以采用其边缘后验概率、对数似然比或者对应接收符号值。每次译码迭代



得到的码符号的可靠性量度的计算结果作为下一次迭代的输入。译码迭代持续进行，直到满足某个特定的停止条件。然后，基于码符号的可靠性量度的计算结果做出判决。迭代概率译码方法属于软判决译码，在译码理论的研究中，根据对接收信号处理方式的不同，译码分为硬判决译码和软判决译码。硬判决译码利用码的代数结构进行译码，比较简单，易于工程实现，但是与软判决译码相比，它在性能上要损失约 2～3dB。软判决译码充分利用了输出波形信息，比硬判决译码具有更大的编码增益。在加性高斯白噪声信道中，它比硬判决译码要多 2dB 的软判决增益，而在衰落信道中，软判决增益超过 5dB。

软判决译码具有两个基本的译码准则，一个是使译码码字错误概率(误字率)最小的逐字软判决译码，另一个是使每个码元误比特率最小的逐位软判决译码。在这两个译码准则意义下的译码算法，它们的性能实际上是一致的，这两个译码准则的重要意义在于允许人们用不同的方法实现软判决译码，也可以从不同角度分析软判决译码。

LDPC 码的优点：具有较大的灵活性和较低的错误平层特性(error floor)；不需要深度交织以获得好的误码性能；描述简单，对严格理论分析具有可验证性；译码不基于网格，复杂度低于 Turbo 码，且可实现完全的并行操作，硬件复杂度低，因而适合硬件实现；吞吐量大，极具高速译码潜力。因此，LDPC 码已成为当前通信编码领域的又一个研究热点。目前在欧洲卫星广播系统 DVB-S52、深空通信、磁记录系统、第四代移动通信的信道编码等领域已准备将 LDPC 码作为信道编码方案之一。另外 LDPC 码与 MIMO、OFDM 等技术的结合也是目前学者们热点研究的问题。

3.8　网 格 编 码

3.8.1　网格编码的基本概念

网格编码调制(Trellis Coded Modulation，TCM)是由 Ungerboeck 在 1982 年提出的一种高级的编码调制方法，这种技术建立在 Ungerboeck 提出的子集划分方法的基础上。TCM 方法的本质是将编码和调制相结合，利用信号集空间的冗余度，以进一步降低误码率。TCM 技术与常规的非编码多进制调制相比具有较大的编码增益且不降低频带利用率，所以特别适合限带信道的信号传输。

网格编码调制有两个基本特征。

(1) 在信号空间中的信号点数目比无编码的调制情况下对应的信号点数目要多(通常增加一倍)，这些增加的信号点使编码有了冗余，而不牺牲带宽。

(2) 采用卷积码编码规则，在一系列信号点之间引入依赖关系，使得只有某些信号点图样是许可使用的信号序列，并可模型化为格状网络，因而称为“网格”编码调制。

3.8.2　网格编码的编译码原理

在传统的数字通信系统中，发送端的纠错编码和调制是分开进行的，接收端的译码和解调也是分开的。纠错编码需要冗余度，在码流中增加监督比特可达到检错或纠错的目的，但此时码流的比特速率将增加，编码后信号速率增加，相应传输带宽需求增加，如果系统

带宽一定，则编码增益需要依靠降低信息传输速率来获得。若要求不降低信息速率，编码增益又需要依靠牺牲带宽来获得。如果系统频带受限，则需要通过增加信号点集以降低因编码冗余引起的带宽增加。信号点集的增加使信号空间距离减小，因此造成误码性能变差，在系统功率受限时，无法保证误码性能。

网格编码是将调制和编码结合起来使用，在加性高斯白噪声信道，这样处理以后就使得决定系统性能的主要参数由卷积码的汉明距离转化为传输信号间的欧氏距离。因此，最佳网格编码的设计是基于欧氏距离，这就要求必须将编码器和调制器当作一个统一的整体进行综合设计，使得编码器和调制器级联后产生的编码信号序列具有最大的欧氏距离。从信号空间的角度看，这种最佳编码调制的设计实际上是一种对信号空间的最佳分割，这就是网格编码调制的基本思想。

TCM 最优码是按照编码信号的网格图确定的，TCM 最优码方案通过一种特殊的信号映射可变成卷积码形式，这种映射的原理是将调制信号集合分割成子集，使得子集内信号间具有更大的欧氏距离。因此 TCM 设计的一个主要目标是寻找与各种调制方式对应的卷积码，当卷积码的每个分支与信号点映射后，使得每条信号路径之间有最大的欧氏距离。根据这个目标，对于多电平/多相位的二维信号空间，把信号点集不断地扩大为 2、4、8、……个子集，使它们之间的信号点的最小距离不断增大，这种映射关系称为集合分割映射。集合分割映射的每一次分割都是将一个较大的信号子集分割成较小的两个子集，这样可得到一个表示集合分割的二叉树，每经过一次分割，子集数加倍，而子集内信号最小距离增加，一直分割到子集内只含欧氏距离最大的两个点。

下面举例说明集合分割映射的具体实现。图 3.22 给出了 8PSK 信号的集合分割映射过程示意图。图中，所有 8 个信号点分布在一个圆周上，经连续 3 次划分后，分别产生 2、4、8 个子集，它们的共同特点是最小码距逐次增大，即 $d_0 < d_1 < d_2$。设最上层星座图圆的半径为 1，则该层信号的最小欧氏距离为 $d_0 = 2\sin(\pi/8) = 0.765$。第一级分割后，得到两个子集，每个子集相当于一个 QPSK，此时信号的最小欧氏距离为 $d_1 = 1.414$。第二级分割后，得到 4 个子集，每个子集相当于一个 2PSK，信号的最小欧氏距离为 $d_2 = 2$。从图 3.22 可以看出每次分割后，信号的欧氏距离都得到增加。

根据集合分割的思想，可以设计比较简单而有效的 TCM 编码方案，TCM 编码调制器的系统结构如图 3.23 所示，设输入码字有 n 个比特，无编码调制时，二维信号空间中应有 2^n 个信号点与之对应。当采用编码调制时，为增加冗余度，有 2^{n+1} 个信号点对应。在图 3.22 中，可划分为 4 个子集，对应于码字的 1 个比特加到编码效率为 1/2 的卷积码输入端，输出 2 个比特，选择相应的子集，码字剩余的未编码数据比特确定信号与子集中信号点之间的映射关系。

TCM 系统使用冗余多进制调制与一个有限状态的网格编码器相结合，由编码器控制选择调制信号，以产生编码符号序列，编码器和调制器级联后产生的编码信号序列具有最大的欧氏自由距离。当编码调制后的信号序列经过一个加性高斯白噪声的信道后，在接收端可以采用维特比算法实现译码，用维特比算法寻找最佳格状路径，以最小码距为准则，执行最大似然序列检测解调出接收的信号序列。在格状图中，每一条支路对应一个子集，



而不是一个信号点。检测的第一步是确定信号点，首先确定信号点所在的子集，在码间距离意义下，这个子集是最靠近接收信号点的子集。信号点确定后，它和接收点之间的平方欧氏距离可用于以后的支路求解，并可用维特比算法继续求解。

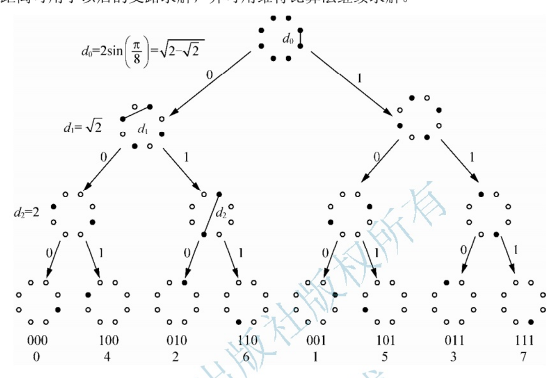


图 3.22　8PSK 信号的集合分割过程

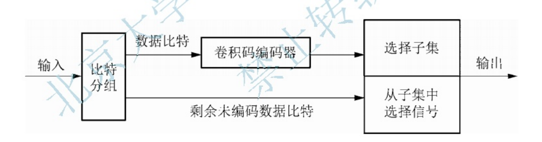


图 3.23　TCM 网格编码系统结构

1988 年 CCITT 将 TCM 技术列为建议 V.33，其基本工作方法就是在话带内以 2400 波特的速率传输 9.6Kbps 的数字语音信号，采用这种技术在电话线上每秒传输 2400 个幅度不等的脉冲，每个脉冲要用 4 位编码表示不同的幅度和相位并加有纠错码，保证正确恢复幅度和相位。编码增益更高的多维 TCM 编码已经用于电视电话，如美国的 AT&T 2500 型电视电话，总传输速率为 19.2Kbps，采用 4 位 16 状态 TCM 编码调制方式，每波特采用 6 位码，传输速率为 3200 波特，实现在一路二线电话线上全双工通信。TCM 技术在各类高速数字传输系统中也推广应用，在卫星通信中，国际卫星组织将利用 TCM 的功率和频谱利用能力强这一优点，在第八代数字卫星系统中即将采用高速 TCM-8PSK 技术，以扩大通信能力。在 IEEE 802.15 和 IEEE 802.16 协议中均采用了 TCM 技术，电力系统数字载波通信中也广泛采用 128TCM 数字调制技术，实现在 8kHz 信道带宽条件下传输速率达到 32Kbps，且比同等信道条件下采用 64QAM 抗噪性能有所改善。

本 章 小 结

本章主要介绍了无线通信编码技术，包括信源编码和信道编码。信源编码的主要目的是提高数据信号传输的有效性，而信道编码的目的是提高数据信号传输的可靠性。信源编码主要是针对语音编码和数据压缩编码，主要信源编码技术有香农编码、霍夫曼编码、游程编码、算术编码、预测编码、变换编码等。

信道编码又称为差错控制编码，它的基本思想是在被传送的信息中加入一些监督码元，通过加入冗余码来减少误码，在收和发之间建立某种检验关系，在接收端实现检错和纠错。常用的差错控制方式主要有前向纠错(FEC)、检错重发(ARQ)、混合纠错(HEC)和信息反馈(IRQ)。主要的信源编码技术有线性分组码、循环码、卷积码、Turbo 码、LDPC 码、TCM 码等。

习　　题

3.1 填空题

(1) 常用的差错控制方式主要有_________、_________、_________和_________。

(2) 线性分组码的主要性质有_________、_________和_________。

(3) 循环码具有两个重要的数学特征：_________和_________。

(4) 汉明码的码字长度 n 与监督码元个数 r 之间的关系为_________，码字长度为 31 的汉明码码字中信息码元个数为_________，此码能纠正发生在一个码字中的_________位错误。

(5) 已知(7,3)循环码的生成多项式为 $g(x)=x^4+x^3+x^2+1$，若输入信息为 110，则系统的输出码字为_________。

3.2 语音编码和数据压缩编码的实现方法有哪些？

3.3 信道编码有哪些分类？为什么要采用差错控制？

3.4 最小码距与其检错和纠错能力有什么关系？

3.5 简述卷积编码、Viterbi 译码的原理。

3.6 什么是 TCM 码？它有哪些特点？其编码原理是如何实现的？

3.7 简述 LDPC 码的编译码原理。

3.8 设信源共有 8 个信源符号，其概率分布为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ 0.37 & 0.16 & 0.14 & 0.13 & 0.07 & 0.06 & 0.04 & 0.03 \end{bmatrix}$$

试对该信源进行香农编码，并计算平均码长及编码效率。

3.9 设单符号离散无记忆信源为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ 0.4 & 0.18 & 0.10 & 0.10 & 0.07 & 0.06 & 0.05 & 0.04 \end{bmatrix}$$



试对该信源进行二进制霍夫曼编码，并计算平均码长及编码效率。

3.10 求下列二元码字之间的汉明距离：

(1) 0000，0101；(2)01110，11100；(3)010101，101001；(4)1110111，1101011。

3.11 某码字的集合为

0000000　　1000111　　0101011　　0011101

1101100　　1011010　　0110110　　1110001

试求：(1)该码字集合的最小码距；(2)根据最小码距确定其检错和纠错能力。

3.12 一个码长 $n=15$ 的汉明码，其监督位 r 应为多少？编码效率是多少？试写出监督码元与信息码元之间的关系。

3.13 已知(7,3)线性分组码的生成矩阵如下，列出所有的许用码组，并求监督矩阵。

$$\boldsymbol{G}=\begin{bmatrix}1&0&0&1&1&1&0\\0&1&0&0&1&1&1\\0&0&1&1&1&0&1\end{bmatrix}$$

3.14 已知(7,3)分组码的监督关系式为

$$\begin{cases}x_6 & & & +x_3 & +x_2 & +x_1 & & = & 0\\ x_6 & & & & +x_2 & +x_1 & +x_0 & = & 0\\ x_6 & +x_5 & & & & +x_1 & & = & 0\\ x_6 & & +x_4 & & & & +x_0 & = & 0\end{cases}$$

求其监督矩阵、生产矩阵、全部码字及纠错能力。

3.15 令 $g(x)=x^3+x+1$ 为(7,4)循环码的生成多项式，(1)求该循环码的生成矩阵和监督矩阵；(2)若两个信息码组分别为 1001 和 0110，求这两个循环码组。

3.16 已知(15,7)循环码，生成多项式 $g(x)=x^8+x^7+x^6+x^4+1$，(1)写出该循环码的生成矩阵；(2)若信息码组为 0011001，写出系统输出码组。

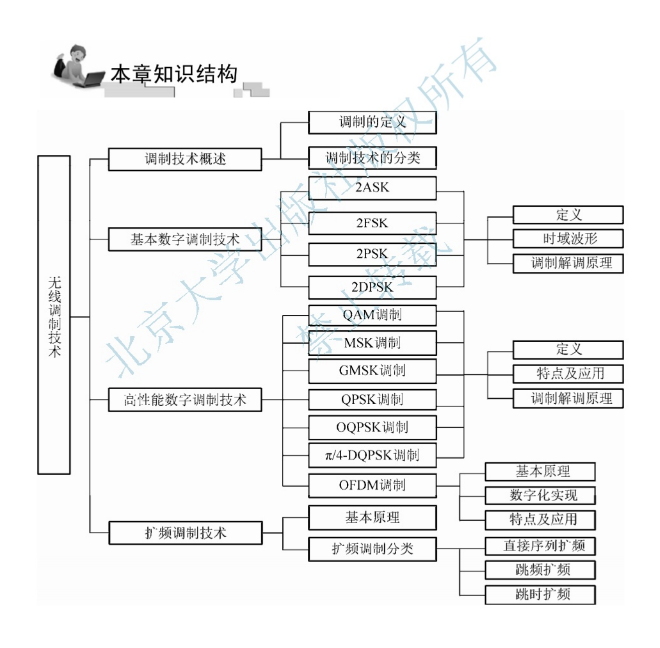
3.17 已知 $x^7+1=(x^3+x^2+1)(x^3+x+1)(x+1)$，(1)写出(7,3)循环码的生成多项式；(2)写出(7,4)循环码的生成多项式；(3)写出(7,6)偶校验码的生成多项式。

3.18 已知一个(3,1,2)卷积码 $g_1(x)=x^2+x+1$，$g_2(x)=x^2+x+1$，$g_3(x)=x^2+1$，则(1)画出该码编码器框图；(2)画出状态图、树图；(3)求该码的自由距离。

第 4 章

无线调制技术

本章知识结构





本章教学目的与要求

- 掌握调制技术的基本概念和分类。
- 了解和掌握数字基带信号的表示形式和各种数字基带信号波形
- 掌握 2ASK、2FSK、2PSK 和 2DPSK 信号的表达式、时域波形、频谱特性和调制解调原理
- 了解二进制数字调制系统的误码率、频带宽度和对信道特性变化的敏感性等性能
- 掌握 QAM 调制技术的原理、特点和星座图
- 掌握 MSK、GMSK 调制技术的基本原理和特点
- 掌握 QPSK、OQPSK、DQPSK、π/4-DQPSK 调制技术的基本原理和特点
- 掌握 OFDM 调制技术的基本原理、数字化实现方法和特点
- 掌握扩频调制的基本原理和扩频方式



引言

无线通信的基本任务是以自由空间为传输信道，把需要传送的信息(声音、文字或图像)变换成无线电波传送到远方的接收点。显然，无线通信中对信息传输的基本要求是：①希望传送距离远；②能实现多路传输，且各路信号传输时，应互不干扰。为了满足这些基本要求，就需要引入调制技术。调制是通信中一个十分重要的概念，是一种信号处理技术。无论在模拟通信、数字通信、还是数据通信中都扮演着重要角色。那么调制的基本概念是什么，有哪些调制方式，其具体原理又是怎么样的呢，通过本章的学习，你将得到答案。

【案例 4.1】邻频调制器

根据国家政策，到 2015 年我国所有县级以上乡村都要基本完成有线电视数字化，数字电视将全部替换模拟电视，这就需要借助邻频调制器实现数字电视改造。邻频调制器(也称作射频调制器、调频调制器或电视调制器)是前端有线电视机房的主要设备之一，其主要功能是把信号源(可以是数字电视机顶盒、卫星数字电视接收机、DVD 机、电脑、摄像机、电视解调器等 AV 信号源)所提供的视频信号(VIDEO)和音频信号(AUDIO)调制成稳定的高频射频振荡信号，其中视频为调幅调制方式，音频为调频调制方式。根据世界上彩色电视制式的不同，邻频调制器有 PAL 制调制器，NTSC 制调制器，SECAM 制调制器 3 种制式，我国采用的是 PAL-D 制式。OKASI-T8604L 邻频调制器(图 4.1)采用四路调制一体设计，中频采用声表滤波器处理，有效地保证了残留边带特性；采用无源、高选择性单频道滤波器，有效地保证了调制器良好的寄生输出抑制性能；采用频率合成技术，频率稳定度高，频率范围广，适用于 49-863MHz 范围的所有系统和电视频道。邻频调制器 T8604L 广泛应用于有线电视前端，将音视频信号调制成射频信号，输入的音频、视频信号可以通过外部的调节旋钮连续调节，允许输入 0.7～1.4V 视频信号源信号、0.775V±10dB 的音频信号，在正常频偏条件下，邻频调制器也可以在外部连续调节使输出图像正常。



图 4.1　邻频调制器

【案例 4.2】QAM 调制器

QAM 调制器(图 4.2)是有线数字电视前端的主要设备之一，接收来自编码器、复用器、DVB 网关、加扰器、视频服务器等设备的 DVB 传输码流，进行 RS 编码、卷积交织和 QAM 调制等信道处理，提供中频输出或者射频输出，可广泛应用于数字电视、数据广播、视频点播、因特网、会议电视等多种数字宽带应用系统。有线电视数字前端、卫星数字电视广播、地面数字电视、图像监控、视频点播(VOD)、远程教学、会议电视功能特征：完全符合 DVB-C 标准的信道编码，支持全部范围的可变符号率和 QAM 映射，自动码流填充及智能空包删除，空包自动过滤和插入，可本地和远程控制，输入码流率显示，码率数据溢出报警；码流格式：多节目传输流(MPTS)和单节目传输流(SPTS)，支持 PID 过滤和节目重映射(可选)，支持 PCR 重校正，PCR 精度<±100ns，可通过网络控制、监测和设置提供商名称，液晶显示，操作方便灵活，运行稳定，可靠性高。

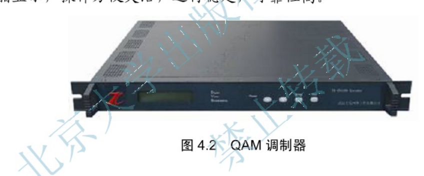


图 4.2　QAM 调制器

4.1　调制技术概述

4.1.1　调制技术的基本概念

调制就是将基带信号加载到高频载波上的过程，其目的是将需要传输的模拟信号或数字信号变换成适合于信道传输的频带信号，以满足无线通信对信息传输的基本要求。如在生活中，我们要将一件货物运到几千千米外的地方，光靠人力本身显然是不现实的，必须借助运载工具来完成，或汽车、或火车、或飞机。如果将运载的货物换成是需要传输的信息，就是通信中的调制了。在这里，货物就相当于调制信号，运载工具相当于载波，将货物装到运载工具上就相当于调制过程，从运载工具上卸下货物就是解调。

调制器的模型如图 4.3 所示，它可以看作一个非线性网络，其中 $m(t)$ 为基带信号，$c(t)$ 为载波，$s_m(t)$ 为已调信号。基带信号是需传送的原始信息的电信号，它属于低频范围。基带信号直接发送存在两个缺点：很难实现多路远距离通信；要求有很长的天线，在工艺及使用上都是很困难的。载波信号是频率较高的高频、超高频甚至微波，若采用无线电发射，

天线尺寸可以很小，并且对于不同的电台，可以采用不同的载波频率，这样接收时很容易区分，就能实现多路互不干扰的传输。

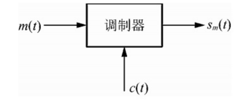


图 4.3　调制器的模型

调制的实质是频谱搬移，即将携带信息的基带信号的频谱搬移到较高的频率范围，基带信号也称调制信号，经过调制后的信号称为频带信号或已调信号。已调信号具有 3 个基本特征：一是携带原始信息，二是适合于信道传输，三是信号的频谱具有带通形式，且中心频率远离零频。

4.1.2　调制技术的分类

根据基带信号 $m(t)$ 、载波 $c(t)$ 和调制器功能的不同，调制技术可以有不同的分类。

根据基带信号 $m(t)$ 形式的不同，调制可分为模拟调制和数字调制。模拟调制中 $m(t)$ 是连续变化的模拟信号，通常以正弦波为代表。常用的模拟调制包括幅度调制(如 AM、DSB、SSB、VSB)、频率调制(FM)和相位调制(PM)。数字调制中 $m(t)$ 是离散变化的数字信号，通常以二进制数字脉冲为代表。常用的数字调制方式有数字振幅调制(ASK)、数字频率调制(FSK)、数字相位调制(PSK)、正交幅度调制(QAM)等。模拟调制和数字调制信号波形举例分别如图 4.4 和图 4.5 所示。

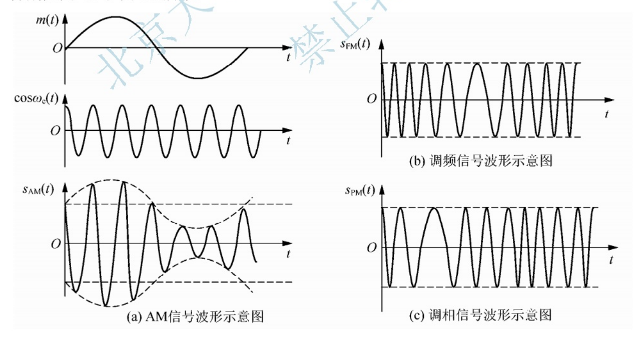
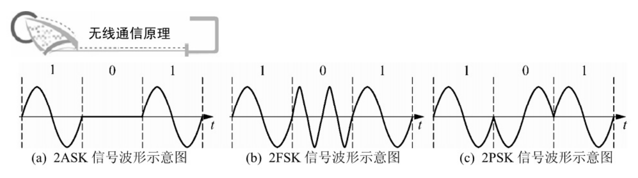


图 4.4　模拟调制信号波形举例



(a) 2ASK 信号波形示意图　(b) 2FSK 信号波形示意图　(c) 2PSK 信号波形示意图

图 4.5　数字调制信号波形举例

根据载波信号 $c(t)$ 形式的不同，调制可分为连续波调制和脉冲调制，其中连续波调制是以正弦型信号作为载波，而脉冲调制是以脉冲序列作为载波。载波信号是完成频谱搬移的工具，由于正弦波具有形式简单、便于产生和接收等特点，因此实际通信中多选用正弦波作为载波信号。

根据基带信号对载波的控制参量(幅度、频率或相位)的不同，调制可分为幅度调制(如 AM、ASK)、频率调制(如 FM、FSK)和相位调制(如 PM、PSK)。幅度调制是指载波的振幅与基带信号的变化呈线性变化关系；频率调制是指载波的频率与基带信号的变化呈线性变化关系；相位调制是指载波的相位与基带信号的变化呈线性变化关系。

根据调制器频谱搬移特性，调制可分为线性调制和非线性调制。线性调制中输出已调信号 $s_m(t)$ 的频谱是基带信号 $m(t)$ 频谱的线性搬移，两者频谱结构相同，如 AM、DSB、SSB、VSB 等幅度调制；非线性调制中输出已调信号 $s_m(t)$ 的频谱与基带信号 $m(t)$ 频谱结构不同，两者不是线性关系，如 FM、PM、FSK、PSK 等调制方式均属于非线性调制。

调制方式的选择对无线通信系统信息传输的有效性和可靠性有着重大的影响。由于数字通信具有抗噪声性能更好，抗信道损耗能力更强，建网灵活，更容易采用数字差错控制技术和数字加密，便于集成化，便于与各种数字终端接口，便于用现代数字信号处理技术对信号进行处理、加工、变换和存储，通信的安全性更好等诸多优点，所以现代无线通信系统大都采用数字调制方式。

本章将重点介绍现代无线通信系统常常使用的各种数字调制技术，如 2ASK、2FSK、2PSK、2DPSK、QPSK、OQPSK、MSK、QAM、OFDM、扩频调制等。

4.2　基本数字调制技术

与模拟调制中的幅度调制、频率调制和相位调制相对应，用数字基带信号改变正弦型载波的幅度、频率或相位中的某个参量，就产生相应的数字振幅调制、数字频率调制和数字相位调制，这 3 种基本方式常常被称为振幅键控(Amplitude Shift Keying，ASK)、移频键控(Frequency Shift Keying，FSK)和移相键控(Phase Shift Keying，PSK)。所谓“键控”是指一种如同“开关”控制的调制方式，“键控”是对数字调制方式的形象描述。基带信号为二进制数字信号时的数字调制方式称为二进制数字调制，在二进制数字调制中，调制信号只有 0 和 1 两个取值，调制后载波的某个参量(幅度、频率或相位)也只能有两个变化状态，这种调制过程就像用调制信号去控制一个开关，从两个具有不同参量的载波中选择相应的载波输出，从而形成已调信号。

下面分别介绍数字基带信号的基本概念以及 ASK、FSK 和 PSK 这 3 种基本的数字调制方式。

4.2.1　数字基带信号

如果数字基带信号各码元波形相同而取值不同，则数字基带信号可表示为

$$s(t)=\sum_{n=-\infty}^{\infty} a_n g\left(t-nT_s\right) \tag{4-1}$$

其中 a_n 是第 n 个码元所对应的电平值，它可以取 0、1 或−1、1 等，T_s 为码元间隔，$g(t)$ 为某种标准脉冲波形，通常为矩形脉冲。

例如：对于二进制数字序列，a_n 取值为 0、1，$g(t)$ 为矩形脉冲，则代码为 1001 的数字基带信号波形如图 4.6 所示。

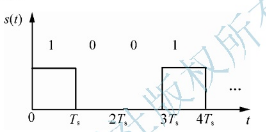


图 4.6　数字基带信号波形举例

常用的数字基带信号波形主要有单极性不归零波形、单极性归零波形、双极性不归零波形、双极性归零波形、差分波形和多电平波形等。

1. 单极性不归零波形

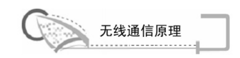
单极性不归零波形用脉冲的零电平和正电平分别对应于二进制的 0 码和 1 码，它是最简单、最常用的一种数字基带信号，其波形举例如图 4.7(a)所示。这种基带信号的特点是极性单一，含有直流分量，在一个码元时间内不是有脉冲就是无脉冲。单极性不归零波形不适合远距离传输，只用于设备内部信号的传输。

2. 双极性不归零波形

双极性不归零波形用脉冲的正、负电平分别对应于二进制的 0 码和 1 码，其波形举例如图 4.7(b)所示。由于该波形是幅度相等极性相反的双极性波形，故当 0、1 等概率出现时无直流分量，此时，恢复信号的最佳判决电平为 0。双极性波形不受信道特性变化的影响，抗干扰能力较强，非常适合于在信道中传输。

3. 单极性归零波形

单极性归零波形中，每个码元有电脉冲时间小于码元时间，每个有电脉冲在小于码元时间内总要回到零电平，如图 4.7(c)所示，这也是单极性归零波形与单极性不归零波形的主要区别。通常用占空比来表示归零波形，设码元持续时间为 T_s，有电脉冲持续时间为 τ，则占空比表示为 τ/T_s，常取 50%，即半占空比。单极性归零波形便于直接提取位定时信息。



4. 双极性归零波形

双极性归零波形是双极性波形的归零形式，即脉冲在每个码元时间内都要回到零电平，相邻脉冲之间总会留有零电位的间隔，其波形举例如图 4.7(d)所示。双极性归零波形具有双极性不归零波形的特点的同时，又有利于同步脉冲的提取。

5. 差分波形

差分波形与上述 4 种波形不同，它的波形不是直接与二进制的 0 码和 1 码本身相对应，而是用相邻脉冲电平的跳变和不跳变来表示二进制的 1 码和 0 码，通常是以电平跳变来表示 1，以电平不跳变来表示 0，其波形举例如图 4.7(e)所示。由于差分波形是以相邻脉冲电平的相对变化来表示代码，因此它又称为相对码波形，而将上述单极性或双极性波形相应地称为绝对码波形。用差分波形作为基带信号传输能够避免出现极性倒置现象，特别是在相位调制系统中用于解决载波相位模糊问题，只要能看出接收波形的电平变化，则接收端均能正确识别。

6. 多电平波形

上述各种信号都是一个二进制码元对应一个脉冲。而多电平波形是多于一个二进制码元对应一个脉冲。图 4.7(f)所示为四电平波形举例，其中每个脉冲对应两个二进制代码。多电平波形在信道带宽一定的条件下，能传输更高的比特率；或在比特率一定时，只需占用更小的传输带宽，因此这种波形在高速数字传输系统中应用广泛。

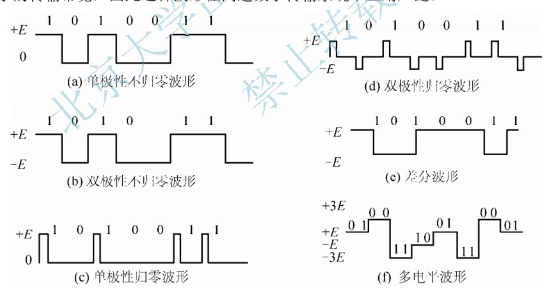


图 4.7 常见的数字基带信号波形

4.2.2 二进制振幅键控

正弦载波的幅度随数字基带信号的变化而变化的数字调制方式称为振幅键控。当载波的振幅随二进制数字基带信号 1 和 0 在两个状态之间变化，而其频率和相位保持不变时，则为二进制振幅键控(2ASK)。设发送的二进制数字基带信号由码元 0 和 1 组成，其中发送



0 码的概率为 P，发送 1 码的概率为 $1-P$，且两者相互独立，则该二进制数字基带信号可表示为

$$s(t)=\sum_{n=-\infty}^{\infty}a_n g\left(t-nT_s\right) \tag{4-2}$$

式中 T_s 为码元间隔，a_n 为符合下列关系的二进制序列的第 n 个码元

$$a_n=\begin{cases}0, & \text{发送概率为 } P \\ 1, & \text{发送概率为 } 1-P\end{cases} \tag{4-3}$$

$g(t)$ 是持续时间为 T_s 的归一化矩形脉冲

$$g(t)=\begin{cases}1, & 0\leqslant t\leqslant T_s \\ 0, & \text{其他}\end{cases} \tag{4-4}$$

则 2ASK 信号的一般时域表达式为

$$s_{2ASK}(t)=\sum_n a_n g\left(t-nT_s\right)\cos\omega_c t \tag{4-5}$$

其中 ω_c 为载波角频率，为了简化，这里假设载波的振幅为 1。由式(4-5)可见，二进制振幅键控(2ASK)信号可以看成是一个单极性矩形脉冲序列与一个正弦型载波相乘。

2ASK 信号的时域波形举例如图 4.8 所示。由图可见，2ASK 信号的波形随二进制基带信号 $s(t)$ 通断变化，因而又被称为通断键控信号(OOK)。2ASK 信号的产生方法有两种：一种是模拟调制法，即按照模拟调制原理来实现数字调制，只需将调制信号由模拟信号改成数字基带信号；另一种是键控调制法，即根据数字基带信号的不同来控制载波信号的“有”和“无”来实现。如当二进制数字基带信号为 1 时，对应有载波输出，当二进制数字基带信号为 0 时，则无载波输出，即载波在数字基带信号 1 或 0 的控制下实现通或断。二进制振幅键控信号的两种产生方法分别如图 4.9 和图 4.10 所示。

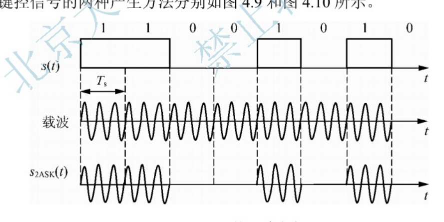


图 4.8　2ASK 信号时域波形

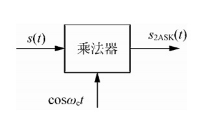


图 4.9　模拟相乘法产生 2ASK 信号

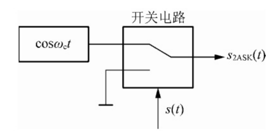


图 4.10　数字键控法产生 2ASK 信号

2ASK 信号的解调可采用非相干解调(包络检波法)和相干解调(同步检测法)方法，两种解调方法的原理框图分别如图 4.11 和图 4.12 所示。2ASK 信号非相干解调过程的时间波形如图 4.13 所示。相干解调需要在接收端接入同频同相的载波，所以又称同步检测。在非相干解调中，全波整流器和低通滤波器构成了包络检波器。

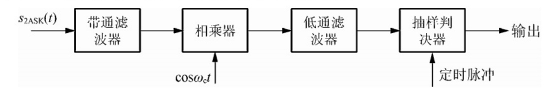


图 4.11　2ASK 信号相干解调原理框图

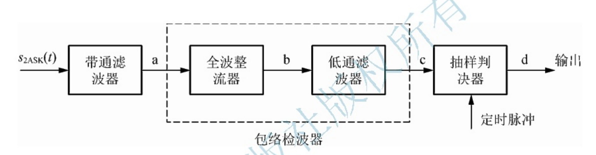


图 4.12　2ASK 信号非相干解调原理框图

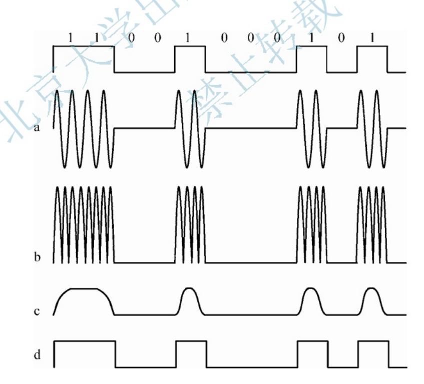


图 4.13　2ASK 信号非相干解调过程的时间波形

2ASK 信号的功率谱密度为数字基带信号功率谱密度的线性搬移，数字基带信号的功率谱密度为 $P_s(f)$，则 2ASK 信号功率谱密度为



$$P_{2\text{ASK}}(f)=\frac{1}{4}\left[P_{\text{s}}\left(f+f_{\text{c}}\right)+P_{\text{s}}\left(f-f_{\text{c}}\right)\right] \tag{4-6}$$

2ASK 信号的功率谱密度如图 4.15 所示，图 4.14 是数字基带信号的功率谱密度，对比两图可知，2ASK 信号频谱处于载频附近，是数字基带信号频谱的线性搬移，并占据一定的带宽。从图 4.15 中可见，2ASK 信号的功率谱包含连续谱和离散谱，其中，连续谱是数字基带信号 $s(t)$ 经线性调制后的双边带频谱，而离散频谱为载波分量，出现在 $\pm f_{\text{c}}$ 处。另外，2ASK 信号的频带宽度 $B_{2\text{ASK}}$ 是数字基带信号带宽的 2 倍。

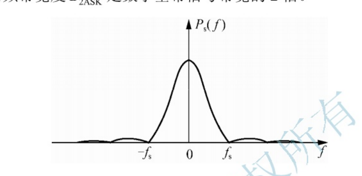


图 4.14　数字基带信号的功率谱密度

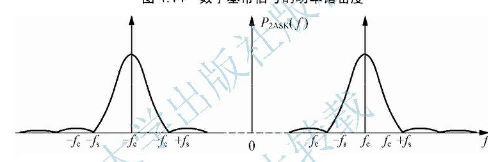


图 4.15　2ASK 信号的功率谱密度

4.2.3　二进制移频键控

移频键控是利用正弦载波的频率变化来表示数字信息，而载波的幅度和初始相位保持不变，如果正弦载波的频率随二进制基带信号 1 和 0 在 f_1 和 f_2 两个频率点间变化，则为二进制移频键控(2FSK)。设发送 1 码时，载波频率为 f_1，发送 0 码时，载波频率为 f_2，则 2FSK 信号的时域表达式为

$$s_{2\text{FSK}}(t)=\left[\sum_n a_n g\left(t-nT_{\text{s}}\right)\right]\cos\omega_1 t+\left[\sum_n \overline{a}_n g\left(t-nT_{\text{s}}\right)\right]\cos\omega_2 t \tag{4-7}$$

其中 $\omega_1=2\pi f_1$， $\omega_2=2\pi f_2$， $\overline{a_n}$ 是 a_n 的取反，即

$$a_n=\begin{cases}0, & \text{概率为}P\\ 1, & \text{概率为}1-P\end{cases} \tag{4-8}$$

$$\overline{a_n}=\begin{cases}1, & \text{概率为}P\\ 0, & \text{概率为}1-P\end{cases} \tag{4-9}$$

从式(4-7)可以看出，2FSK 信号可以看成是两个不同载频交替发送的 2ASK 信号的叠加。2FSK 信号时域波形如图 4.16 所示。

2FSK 信号的产生可以采用模拟调频电路和数字键控两种方法实现。图 4.17 是用数字键控的方法产生二进制移频键控信号的原理图，图中两个振荡器的输出载波受输入的二进制基带信号控制，在一个码元时间 T_s 内只输出 f_1 或 f_2 这两个载波的其中一个。

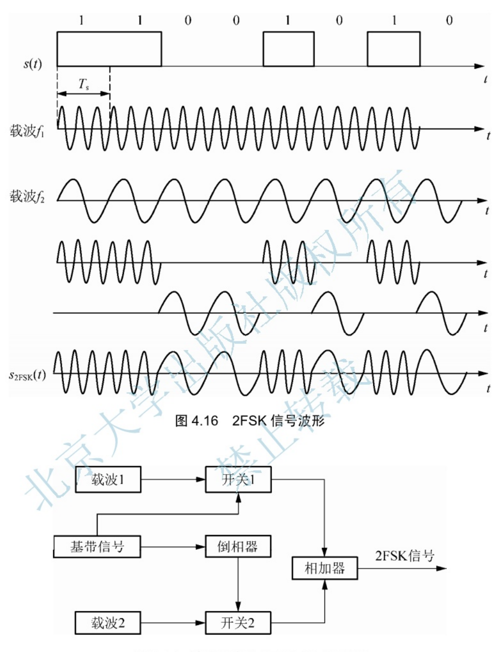


图 4.16　2FSK 信号波形

图 4.17　键控法产生 2FSK 信号原理图

2FSK 信号的解调方法很多，如模拟鉴频法、数字检测法、非相干解调法和相干解调法。2FSK 信号非相干解调和相干解调两种方法的原理图分别如图 4.18 和图 4.19 所示。两种方法的解调过程都是将二进制移频键控信号分解为上下两路二进制振幅键控信号分别进行解调，通过对上下两路的抽样值进行比较最终判决得到输出信号。2FSK 信号非相干解调过程的时间波形如图 4.20 所示。

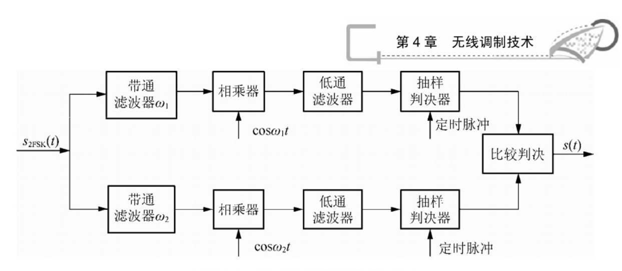


图 4.18　2FSK 信号相干解调原理图

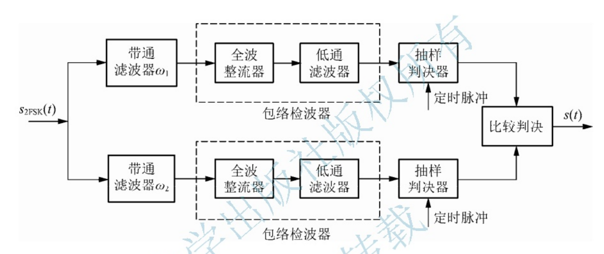


图 4.19　2FSK 信号非相干解调原理图

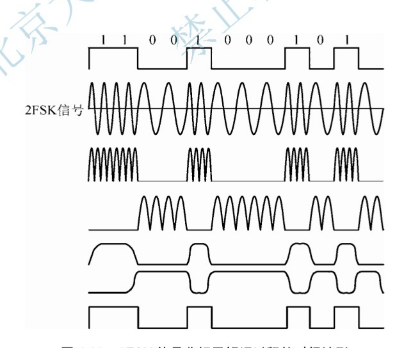


图 4.20　2FSK 信号非相干解调过程的时间波形



2FSK 信号的功率谱密度可用两个中心频率分别为 f_1 和 f_2 的 2ASK 信号的功率谱来表示，即

$$P_{2FSK}(f)=\frac{1}{4}\left[P_s(f-f_1)+P_s(f+f_1)+P_s(f-f_2)+P_s(f+f_2)\right] \tag{4-10}$$

2FSK 信号的功率谱如图 4.21 所示，2FSK 信号的功率谱与 2ASK 信号的功率谱相似，同样包含连续谱和离散谱。2FSK 信号中，连续谱由两个中心频率位于 f_1 和 f_2 处的双边谱叠加而成，离散谱位于两个载频 f_1 和 f_2 处。对于 2FSK 信号，通常可定义其移频键控指数

$$h=|f_1-f_2|/f_s \tag{4-11}$$

显然，不同的 h，与之对应的 2FSK 信号的功率谱是不一样的。当 $h<1$ 时，2FSK 信号的功率谱与 2ASK 的极为相似，连续谱在 $f_c=(f_1+f_2)/2$ 处呈单峰状；当 $h>1$ 时，2FSK 信号的功率谱呈双峰状。若以 2FSK 信号功率谱第一个零点之间的频率间隔定义为二进制移频键控信号的带宽，则该二进制移频键控信号的带宽 B_{2FSK} 为

$$B_{2FSK}=|f_1-f_2|+2f_s \tag{4-12}$$

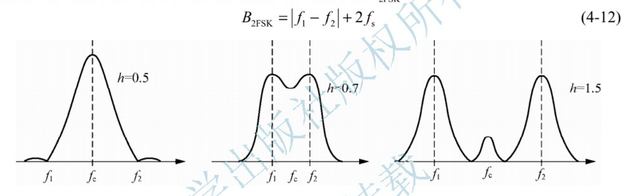


图 4.21 2FSK 信号的频谱(两个频率差对功率谱的影响)

4.2.4 二进制移相键控

1. 2PSK 调制

移相键控是指正弦载波的相位随数字基带信号离散变化，二进制移相键控(2PSK)是用二进制数字基带信号控制载波的相位变化有两个状态，例如，二进制数字基带信号的 1 和 0 分别对应着载波的相位 0 和 π。

二进制移相键控信号表达式为

$$s_{2PSK}(t)=\left[\sum_n a_n g(t-nT_s)\right]\cos\omega_c t \tag{4-13}$$

其中 a_n 为双极性数字信号，即

$$a_n=\begin{cases}+1, & 概率为 P\\ -1, & 概率为 1-P\end{cases} \tag{4-14}$$

当 $g(t)$ 是持续时间为 T_s 的归一化矩形脉冲时，有

$$s_{2PSK}(t)=\begin{cases}\cos\omega_c t, & 概率为 P\\ -\cos\omega_c t=\cos(\omega_c t+\pi), & 概率为 1-P\end{cases} \tag{4-15}$$

由式(4-15)可见，当发送 1 时，2PSK 信号载波相位为 0，发送 0 时载波相位为 π，若用 φ_n

表示第 n 个码元的相位，则有

$$\varphi_n=\begin{cases}0, & \text{发送“1”}\\ \pi, & \text{发送“0”}\end{cases} \tag{4-16}$$

这种二进制数字基带信号直接与载波的不同相位相对应的调制方式通常称为二进制绝对移相调制。2PSK 信号时域波形如图 4.22 所示。

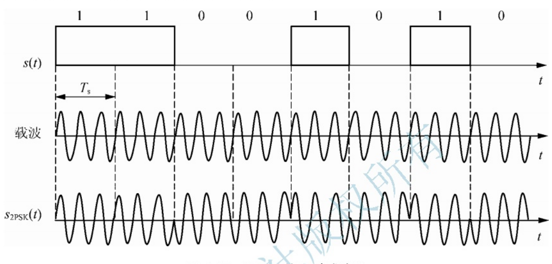


图 4.22　2PSK 信号时域波形

2PSK 信号的产生可以采用模拟调制和数字键控两种方法实现。图 4.23 所示为模拟调制法生成 2PSK 信号，二进制数字序列经码型变换，由单极性码形成双极性不归零码，与载波相乘而产生 2PSK 信号。图 4.24 为数字键控法生成 2PSK 信号，数字键控法中仍用单极性基带信号控制开关电路，当 $s(t)$ 为 1 时，开关电路连接 $\cos\omega_c t$ 振荡器，持续时间 T_s；而当 $s(t)$ 为 0 时，开关电路连接振荡信号的 180° 移相输出，持续时间仍为 T_s，由此得到 2PSK 信号。

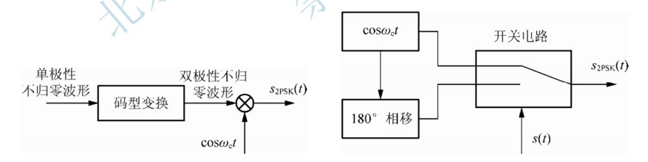


图 4.23　模拟调制法产生 2PSK 信号　　图 4.24　键控法产生 2PSK 信号

2PSK 信号进行解调通常采用相干解调方式，其解调器原理框图如图 4.25 所示。2PSK 信号相干解调各点时间波形如图 4.26 所示。2PSK 信号是用载波的不同相位直接表示相应的数字信号，在这种绝对移相的方式中，由于发送端是以某一个相位作为基准的，因而在接收系统中也必须有这样一个固定基准相位作参考。当恢复的相干载波产生 180°倒相时，其解调恢复的数字信息就会与发送的数字信息完全相反，从而造成解调器输出数字基带信号全部出错，这种现象称为 2PSK 的随机“倒π”现象。

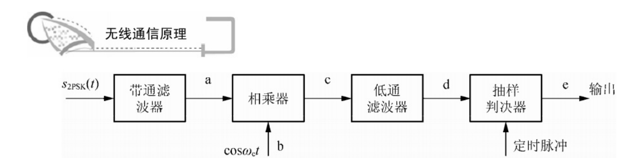


图 4.25　2PSK 信号相干解调原理框图

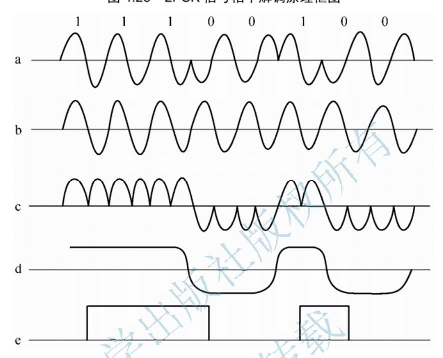


图 4.26　2PSK 信号相干解调各点波形示意图

2PSK 信号可表示为双极性不归零二进制基带信号与正弦载波相乘，则 2PSK 信号的功率谱密度可以看成是双极性基带信号的功率谱密度在 $\pm f_c$ 频率处的频谱搬移，其一般表达式为

$$P_{2PSK}(f)=\frac{1}{4}\left[P_s(f-f_c)+P_s(f+f_c)\right] \tag{4-17}$$

2PSK 信号的功率谱密度如图 4.27 所示。一般情况下 2PSK 信号的功率谱密度由离散谱和连续谱组成，其结构与 2ASK 信号的功率谱密度相类似，带宽也是基带信号带宽的两倍。当二进制基带信号的 1 码和 0 码出现概率相等时，则不存在离散谱。

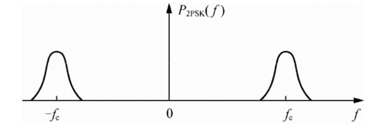
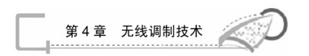


图 4.27　2PSK 信号的功率谱密度



2. 2DPSK 调制

由于 2PSK 调制方式在解调过程中会产生随机“倒π”现象，所以实际中一般不采用 2PSK 方式，而常采用二进制差分移相键控(2DPSK)。2DPSK 调制方式是用前后相邻码元的载波相对相位变化来表示数字信息，所以又称为相对移相键控。假设$\Delta\varphi$为前后相邻码元的载波相位差，可定义一种数字信息与$\Delta\varphi$之间的关系

$$\Delta\varphi=\begin{cases}0, & \text{表示数字信息“0”}\\ \pi, & \text{表示数字信息“1”}\end{cases} \tag{4-18}$$

或

$$\Delta\varphi=\begin{cases}\pi, & \text{表示数字信息“0”}\\ 0, & \text{表示数字信息“1”}\end{cases} \tag{4-19}$$

则数字信息序列与 2DPSK 信号的相位关系可举例表示为

数字信息：		0	0	1	1	1	0	0	1	0	1
2DPSK 信号相位：	0	0	0	π	0	π	π	π	0	0	π
或	π	π	π	0	π	0	0	0	π	π	0

2DPSK 信号调制器原理图如图 4.28 所示。2DPSK 信号的实现过程：首先对二进制数字基带信号进行差分编码，将绝对码表示二进制信息变换为用相对码表示二进制信息，然后再根据相对码进行绝对调相，从而产生 2DPSK 信号。2DPSK 信号调制过程波形如图 4.29 所示。

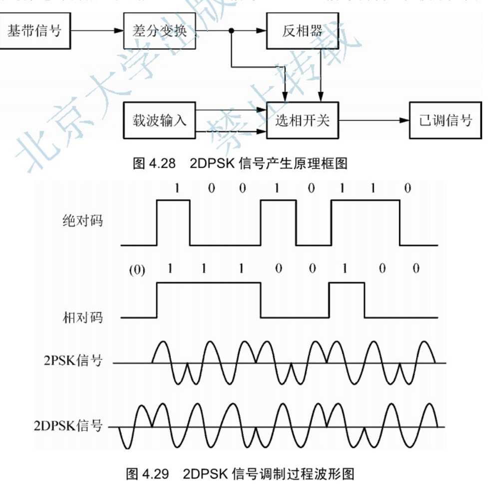


图 4.28　2DPSK 信号产生原理框图

图 4.29　2DPSK 信号调制过程波形图

2DPSK 信号可以采用相干解调方式进行解调，其解调过程是先对 2DPSK 信号进行相干解调，恢复出相对码，再通过码反变换器将相对码变换为绝对码，从而恢复出发送的二进制数字信息。解调器原理图和解调过程各点时间波形分别如图 4.30 和图 4.31 所示。2DPSK 波形的同一相位并不对应相同的数字信息，而前后码元相对相位的差才唯一决定数字信息，这说明，解调 2DPSK 信号时并不依赖于某一固定的载波相位参考值，只要前后码元的相对相位关系不破坏，则只要鉴别这个关系就可以正确恢复数字信息，这就避免了 2PSK 方式中的“倒π”现象发生。此外，单纯从波形上看，2PSK 与 2DPSK 信号是无法分辨的。这说明，一方面，只有在已知移相键控方式是绝对的还是相对的前提下，才能正确判定原信息；另一方面，相对移相信号可以看成是将数字信息(绝对码)变换成相对码(差分码)，然后再根据相对码进行绝对移相而形成的。2DPSK 相干解调与 2PSK 相干解调是相似的，区别仅在于 2DPSK 相干解调系统中有一个码型反变换模块，其作用是进行差分译码，这与调制端的差分编码是对应的。

2DPSK 信号的功率谱密度与 2PSK 信号的功率谱密度是相同的，这里不再赘述。

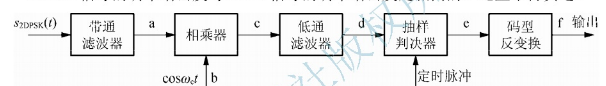


图 4.30 2DPSK 解调原理框图

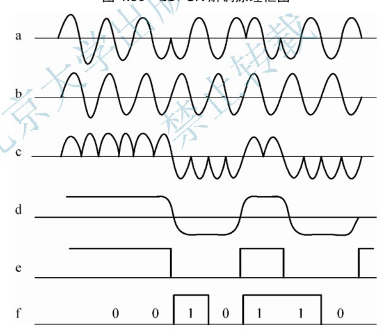


图 4.31 2DPSK 信号解调过程各点波形

4.2.5 二进制数字调制系统性能比较

下面将对 2ASK、2FSK、2DPSK 这 3 种二进制数字调制系统的性能进行比较。

1. 误码率

对于 2ASK、2FSK、2PSK 及 2DPSK，每种数字调制方式又有相干解调方式和非相干解调方式。表 4-1 列出了各种二进制数字调制系统的误码率与输入信噪比 r 的数学关系，其中 $r = a^2/2\sigma^2$ ，a 为已调信号振幅，σ^2 为噪声方差。

表 4-1　二进制数字调制系统的误码率比较

调制方式	误码率	
	相干解调	非相干解调
2ASK	$\mathrm{erfc}\sqrt{\frac{r}{4}}$	$\frac{1}{2}\exp\left(-\frac{r}{4}\right)$
2FSK	$\frac{1}{2}\mathrm{erfc}\sqrt{\frac{r}{2}}$	$\frac{1}{2}\exp\left(-\frac{r}{2}\right)$
2PSK/2DPSK	$\frac{1}{2}\mathrm{erfc}\sqrt{r}$	$p_e = \frac{1}{2}\exp(-r)$

从表 4-1 可看出，对同一种数字调制方式，相干解调方式略优于非相干解调方式。3 种相干(或非相干)解调方式之间，在相同误码率条件下，2ASK、2FSK 和 2PSK 这 3 种系统所需要的信噪比关系为

$$r_{2\mathrm{ASK}} = 2r_{2\mathrm{FSK}} = 4r_{2\mathrm{PSK}} \tag{4-20}$$

将上式转换为分贝表示形式

$$\left(r_{2\mathrm{ASK}}\right)_{\mathrm{dB}} = \left(r_{2\mathrm{FSK}}\right)_{\mathrm{dB}} + 3\mathrm{dB} = \left(r_{2\mathrm{PSK}}\right)_{\mathrm{dB}} + 6\mathrm{dB} \tag{4-21}$$

可见，在相同误码率条件下，在信噪比要求上 2PSK 比 2FSK 小 3dB、2FSK 比 2ASK 小 3dB。由此看来，在抗加性高斯白噪声方面，相干 2PSK 性能最好，2FSK 次之，2ASK 最差。反过来，若信噪比 r 一定，2PSK 系统的误码率低于 2FSK 系统，2FSK 系统的误码率低于 2ASK 系统。对 2ASK、2FSK、2PSK 来说，发送端设备的复杂程度差别不大，而接收端的复杂程度则与所选用的调制和解调方式有关。对于同一种调制方式，相干解调的设备要比非相干解调的复杂；而同为相干解调时，2PSK 的设备最复杂，2FSK 次之，2ASK 最简单。

2. 频带宽度

当码元宽度为 T_s 时，2ASK 系统和 2PSK 系统具有相同的频带宽度，它们的频带宽度都近似为 $2/T_s$ ，2FSK 系统的频带宽度近似为 $|f_2 - f_1| + 2/T_s$ 。可见，2FSK 系统的频带宽度要大于 2ASK 系统或 2PSK 系统的频带宽度，从频带利用率上看，2FSK 系统的频带利用率最低。

3. 对信道特性变化的敏感性

在选择数字调制方式时，还应考虑系统对信道特性变化的敏感程度。在 2FSK 系统中，不需要人为设置判决门限，判决器直接比较两路解调输出的大小作出判断，因而 2FSK 系

统对信道变化较不敏感。在 2PSK 系统中，判决器的最佳判决门限电平为零，与接收机输出信号的幅度无关，判决门限不随信道特性的变化而变化。对于 2ASK 系统，当 $P(0)=P(1)$ 时，判决器最佳判决门限为 $a/2$，它与接收机输入信号的幅度有关。接收机输入信号的幅度 a 将随信道特性的变化而发生变化，相应地，判决器的最佳判决门限也将随之改变，这时，接收机不容易保持在最佳判决门限状态，从而导致误码率增大。因此，2ASK 系统对信道特性变化较敏感，性能最差。

当信道存在严重的衰落时，通常采用非相干解调，因为这时在接收端不易得到相干解调所需的相干载波。当发射机有严格的限制时，可考虑采用相干解调。因为在给定的码元传输速率及误码率的条件下，相干解调所要求的信噪比要比非相干解调小。在恒参信道传输中，如果要求较高的功率利用率，则应选择相干 2PSK 和 2DPSK，而 2ASK 最不可取。如果要求较高的频带利用率，则应选择相干 2PSK 和 2DPSK，而 2FSK 最不可取。

4.3 高性能数字调制技术

2ASK、2FSK 和 2PSK 这 3 种基本的数字调制技术存在频谱利用率较低、抗多径衰落性能较差、功率谱衰减慢、严重带外辐射等诸多缺点。因此为了满足各种无线通信数据传输的要求，人们不断提出新的抗干扰性强、误码率低、频谱利用率高的高性能数字调制技术。正交幅度调制(QAM)和正交频分复用(OFDM)调制方式具有较高的频谱利用率，高斯最小移频键控(Gaussian Filtered Minimum Shift Keying，GMSK)和π/4 正交移相键控(Differential Quadrature Reference Phase Shift Keying，π/4 -DQPSK)具有带外辐射小、抗多径衰落性能好等优点，其他一些新型的数字调制方式如最小移频键控(Minimum Shift Keying，MSK)，四相移相键控(Quadrature Reference Phase Shift Keying，QPSK)，交错正交四相移相键控(Offset Quadrature Reference Phase Shift Keying，OQPSK)，四相相对移相键控(Differential Quadrature Reference Phase Shift Keying，DQPSK)，也已在数字蜂窝移动通信系统中得到广泛应用。本节将主要介绍这几种高性能的数字调制技术。

4.3.1 正交振幅调制

正交振幅调制(Quadrature Amplitude Modulation，QAM)是一种矢量调制，它是幅度和相位联合的调制技术。QAM 是用两路独立的数字基带信号对两个相互正交的同频载波进行抑制载波的双边带调制，利用已调信号在同一带宽内频谱正交的性质来实现两路并行的数字信息传输。由于调制信号是通过电平幅度和载波相位承载比特信息，不同的幅度和相位代表不同的编码，因此，QAM 可以支持更高的调制阶数，而且抗噪声能力强，实现技术简单。QAM 调制的频谱利用率很高，是目前大容量数字微波通信系统、有线电视网络高速数据传输、卫星通信系统等领域广泛应用的一种数字调制方式。

QAM 信号的一般表达式为

$$s_{\mathrm{MQAM}}(t)=X_k\cos\omega_{\mathrm{c}}t+Y_k\sin\omega_{\mathrm{c}}t \tag{4-22}$$

其中 X_k、Y_k 为离散的振幅值，由输入数据确定，(X_k, Y_k)决定了已调 QAM 信号在信号空



间中的坐标点，$k=1,2,\cdots,M$，M 为 X_k 和 Y_k 的个数。

矢量端点的分布图称为星座图，通常用星座图来描述 QAM 信号的信号空间分布状态。若 QAM 的同相和正交支路均采用二进制信号，则信号空间中的坐标点数目(状态数) $M=4$，记为 4QAM。若同相和正交支路都采用四进制信号将得到 16QAM 信号。依此类推，两条支路都采用 L 进制信号将得到 MQAM 信号，其中 $M=L^2$。图 4.32 为 MQAM 信号的矩形排列星座图。

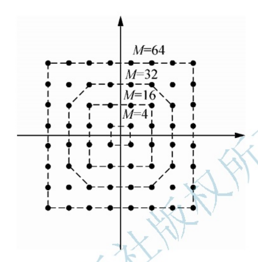


图 4.32　MQAM 信号的星座图

常用 QAM 调制的阶数有 4QAM、16QAM、32QAM 和 64QAM。对于 $M=16$ 的 16QAM 来说，有多种分布形式的信号星座图，其中两种具有代表意义的信号星座图如图 4.33 和图 4.34 所示。在图 4.33 中，信号点的分布成方型，故称为方型 16QAM 星座图，也称为标准型 16QAM。在图 4.34 中，信号点的分布成星型，故称为星型 16QAM 星座图。

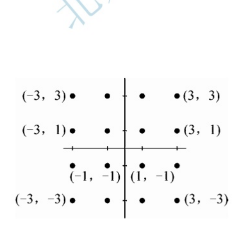


图 4.33　方型 16QAM 星座图

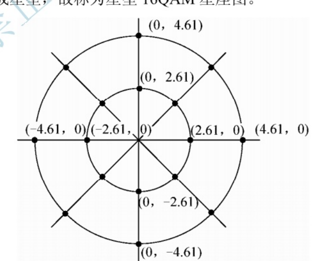


图 4.34　星型 16QAM 星座图

若信号点之间的最小距离为 2，且所有信号点等概率出现，则平均发射信号功率记为

$$P_s=\frac{1}{M}\sum_{k=1}^{M}\left(X_k^2+Y_k^2\right) \tag{4-23}$$

对于方型 16QAM，信号平均功率为

$$P_s = \frac{1}{M}\sum_{k=1}^{M}\left(X_k^2 + Y_k^2\right) = \frac{1}{16}(4\times2 + 8\times10 + 4\times18) = 10$$

对于星型 16QAM，信号平均功率为

$$P_s = \frac{1}{M}\sum_{k=1}^{M}\left(X_k^2 + Y_k^2\right) = \frac{1}{16}\left(8\times2.61^2 + 8\times4.61^2\right) = 14.03$$

由以上两式可以看出，两者信号平均功率是存在差异的，此外，两者的星座结构也有重要的差别，方型星座图中，信号点共有 3 种振幅值和 12 种相位值，而星型星座图中，信号点共有 2 种振幅值和 8 种相位值。在无线移动通信的环境中，存在多径效应和各种干扰，信号振幅和相位的取值种类越多，受到的影响就越大，接收端也就越难恢复原信号，因此在衰落信道中，星型 16QAM 比方型 16QAM 的性能更优。

图 4.35 是 QAM 信号的调制生成框图，其中，输入的二进制序列经过串/并转换器输出速率减半的两路并行序列，分别经过 2 电平到 L 电平的变换，形成 L 电平的基带信号，然后经过预调制低通滤波器分别形成 X_k 和 Y_k，再分别与同相载波和正交载波相乘，最后将两路信号相加即可得到 QAM 信号。

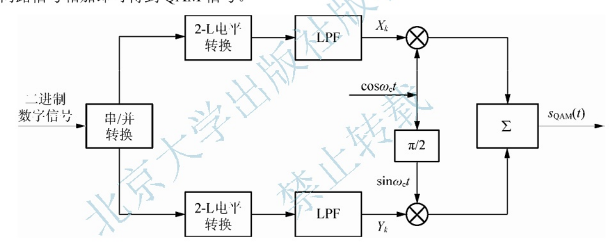
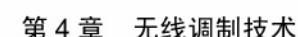


图 4.35　QAM 信号产生框图

图 4.36 是 QAM 信号解调原理框图。QAM 信号同样可以采用正交相干解调方法，解调器输入信号与本地恢复的两个正交载波相乘后，经过低通滤波输出两路多电平基带信号 X_k 和 Y_k，多电平判决器对多电平基带信号进行判决和检测后，再经 L 电平到 2 电平转换和并/串转换器最终输出二进制数字信号。

QAM 调制技术具有能充分利用带宽、抗噪声能力强等优点，有利于实现高速数字传输。由于 QAM 的星座点比 PSK 的星座点更分散，星座点之间的距离因此更大，所以能提供更好的传输性能。但是 QAM 星座点的幅度不是完全相同的，所以它的解调器需要能同时正确检测相位和幅度，不像 PSK 解调只需要检测相位，这增加了 QAM 解调器的复杂性。



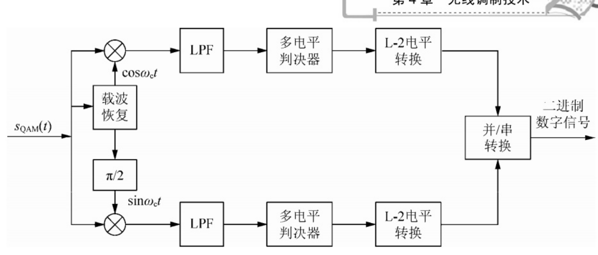


图 4.36　QAM 信号解调框图

4.3.2　最小移频键控

1. MSK 的基本原理

最小移频键控(Minimum Frequency Shift Keying，MSK)是 FSK 的一种改进型，这里“最小”是指能以最小的调制指数(即 0.5)获得正交信号，MSK 比 PSK 传送的比特速率更高，且带外的频谱分量比 2PSK 衰减得更快。

二进制 MSK 信号的表达式可写为

$$s_{\mathrm{MSK}}(t)=A\cos\left[\omega_{c}t+\theta_{k}(t)\right]=A\cos\left(\omega_{c}t+\frac{\pi a_{k}}{2T_{s}}t+\varphi_{k}\right),\quad kT_{s}\leqslant t\leqslant(k+1)T_{s} \tag{4-24}$$

其中 A 为载波振幅，ω_{c} 表示载频，$\pi a_{k}/2T_{s}$ 表示相对载频的频偏，$a_{k}=\pm1$ 是数字基带信号第 k 个码元数据，$\theta_{k}(t)=\dfrac{\pi a_{k}t}{2T_{s}}+\varphi_{k}$ 称为附加相位函数，它表示除载波相位之外的附加相位，φ_{k} 表示第 k 个码元的起始相位。

令

$$\phi_{k}(t)=\omega_{c}t+\frac{\pi a_{k}}{2T_{s}}t+\varphi_{k} \tag{4-25}$$

则

$$\frac{\mathrm{d}\phi_{k}(t)}{\mathrm{d}t}=\omega_{c}+\frac{\pi a_{k}}{2T_{s}}=\begin{cases}\omega_{c}+\dfrac{\pi}{2T_{s}}, & a_{k}=+1\\ \omega_{c}-\dfrac{\pi}{2T_{s}}, & a_{k}=-1\end{cases} \tag{4-26}$$

当 $a_{k}=-1$ 时，信号频率为

$$f_{1}=f_{c}-\frac{1}{4T_{s}} \tag{4-27}$$

当 $a_{k}=+1$ 时，信号频率为

$$f_{2}=f_{c}+\frac{1}{4T_{s}} \tag{4-28}$$

因此可计算出频差

$$\Delta f = f_2 - f_1 = \frac{1}{2T_s} \tag{4-29}$$

即最小频差等于码元传输速率的一半，对应的调制指数为

$$h = \frac{\Delta f}{f_s} = \Delta f \times T_s = \frac{1}{2T_s} \times T_s = 0.5 \tag{4-30}$$

中心频率为

$$f_c = \frac{n}{4T_s},\ n = 1, 2, \cdots \tag{4-31}$$

2. MSK 信号的相位连续性

相位连续是指不仅在一个码元持续期间相位连续，而且在从码元 a_{k-1} 到 a_k 转换的时刻 kT_s，两个码元的相位也连续。对于 MSK 调制，要保证 MSK 信号相位在码元转换时刻连续，则要求对第 k 个码元的相位常数 φ_k 在 $t = kT_s$ 时间内应满足如下约束条件

$$\varphi_k = \varphi_{k-1} + (a_{k-1} - a_k)\frac{\pi(k-1)}{2} = \begin{cases} \varphi_{k-1}, & a_k = a_{k-1} \\ \varphi_{k-1} \pm (k-1)\pi, & a_k \neq a_{k-1} \end{cases} \tag{4-32}$$

为简便起见，设第一个码元的起始相位为 0，则 $\varphi_k = 0$或$\pm\pi$。上式反映了 MSK 信号前后码元区间的相位约束关系，表明 MSK 信号在第 k 个码元的相位常数不仅与当前码元的取值 a_k 有关，而且还与前一码元的取值 a_{k-1} 及相位常数 φ_{k-1} 有关。

下面讨论在每个码元间隔 T_s 内相对于载波相位的附加相位函数的变化。

由 $\theta_k(t) = \dfrac{\pi a_k}{2T_s}t + \varphi_k$ 可知，$\theta_k(t)$ 是 MSK 信号的总相位减去随时间线性增长的载波相位得到的剩余相位，它是一个直线方程式。在一个码元间隔内，当 $a_k = +1$ 时，$\theta_k(t)$ 增大 $\pi/2$；当 $a_k = -1$ 时，$\theta_k(t)$ 减小 $\pi/2$，也就是说相对前一码元附加相位，$\theta_k(t)$ 不是增加 $\pi/2$，就是减小 $\pi/2$。对于各种可能的输入信号序列，$\theta_k(t)$ 随时间的变化规律如图 4.37 所示。它是一个从 -2π 到 2π 的网格图，其中正斜率直线表示传 1 码的相位轨迹，负斜率直线表示传 0 码的相位轨迹。从该网格图中，可以找出任意一个数据序列由零开始的路径。例如，图中黑色加粗线条为数据序列 11101000011 的相位轨迹。

综上，总结 MSK 信号有如下特点。

(1) MSK 信号的包络恒定不变。

(2) MSK 是调制指数为 0.5 的正交信号，频率偏移等于 $\pm 1/4T_s$ Hz。

(3) MSK 波形相位在码元转换时刻是连续的。

(4) MSK 附加相位在一个码元持续时间内线性地变化 $\pm\pi/2$。



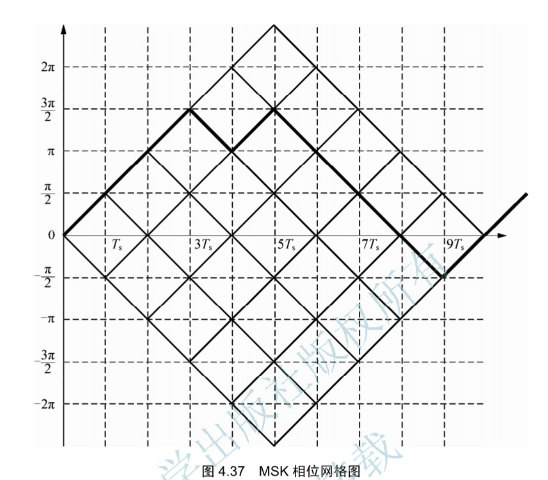


图 4.37　MSK 相位网格图

【例 4.1】已知载波频率 $f_c = 1.75/T_s$，初始相位 $\varphi_0 = 0$。求：(1)当数字基带信号 $a_k = \pm 1$ 时，MSK 信号的两个频率 f_1 和 f_2；(2)对应的最小频差及调制指数；(3)若基带信号为 +1-1-1+1+1+1，画出相应的 MSK 信号波形。

解：(1)当 $a_k = -1$ 时，信号频率 f_1 为

$$f_1 = f_c - \frac{1}{4T_s} = \frac{1.75}{T_s} - \frac{1}{4T_s} = \frac{1.5}{T_s}$$

当 $a_k = +1$ 时，信号频率 f_2 为

$$f_2 = f_c + \frac{1}{4T_s} = \frac{1.75}{T_s} + \frac{1}{4T_s} = \frac{2}{T_s}$$

(2) 最小频差 Δf 为

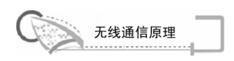
$$\Delta f = f_2 - f_1 = \frac{2}{T_s} - \frac{1.5}{T_s} = \frac{1}{2T_s}$$

它等于码元传递速率的一半。

调制指数为

$$h = \frac{\Delta f}{f_s} = \Delta f \times T_s = \frac{1}{2T_s} \times T_s = 0.5$$

(3) 根据以上计算结果，可以画出相应的 MSK 波形如图 4.38 所示。“+1”和“−1”对



应 MSK 波形相位在码元转换时刻是连续的，而且在一个码元期间所对应的波形恰好相差 1/2 载波周期。

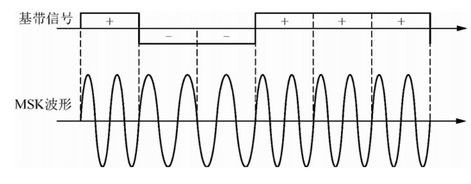


图 4.38　MSK 波形

3. MSK 信号的调制与解调

$$
\begin{aligned}
S_{\text{MSK}}(t) &= \cos\left(\omega_c t + \frac{\pi a_k}{2T_s} t + \varphi_k\right) \\
&= \cos\varphi_k \cos\left(\frac{\pi}{2T_s} t\right)\cos\omega_c t - a_k \cos\varphi_k \sin\left(\frac{\pi}{2T_s} t\right)\sin\omega_c t
\end{aligned}
\tag{4-33}
$$

其中 $(k-1)T_s \leqslant t \leqslant kT_s$，令 $I_k = \cos\varphi_k$，$Q_k = -a_k\cos\varphi_k$，则

$$
S_{\text{MSK}}(t) = I_k \cos\left(\frac{\pi}{2T_s} t\right)\cos\omega_c t + Q_k \sin\left(\frac{\pi}{2T_s} t\right)\sin\omega_c t,\ (k-1)T_s \leqslant t \leqslant kT_s \tag{4-34}
$$

MSK 信号调制器原理图如图 4.39 所示，图中输入二进制数字信号经过差分编码后进行串/并转换，将其中一路延迟 T_s，得到相互交错一个码元宽度的两路信号 I_k 和 Q_k，加权函数 $\cos(\pi t/2T_s)$ 和 $\sin(\pi t/2T_s)$ 分别对两路数据信号 I_k 和 Q_k 进行加权，加权后的两路信号再分别对正交载波 $\cos\omega_c t$ 和 $\sin\omega_c t$ 进行调制，调制后的信号相加后通过带通滤波器，就得到 MSK 信号。

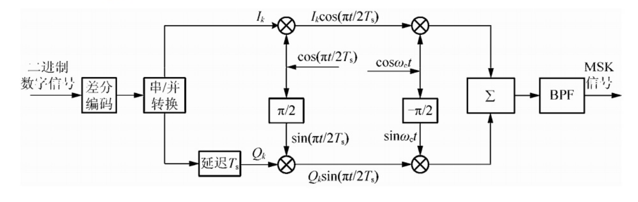


图 4.39　MSK 信号产生原理框图

MSK 信号可以采用相干解调和非相干解调。图 4.40 是 MSK 信号相干解调器原理图，MSK 信号经带通滤波器后分别与相互正交的相干载波相乘，从而将 I_k 和 Q_k 两路信号区分开，然后通过低通滤波器后再进行抽样判决，同相支路在 $2kT_s$ 时刻抽样，正交支路在



$(2k+1)T_s$ 时刻抽样，判决器根据抽样后的信号极性进行判决，大于 0 判为"1"，小于 0 判为"0"，经并/串转换，变为串行数据，最后经差分译码即可恢复原数字信息。

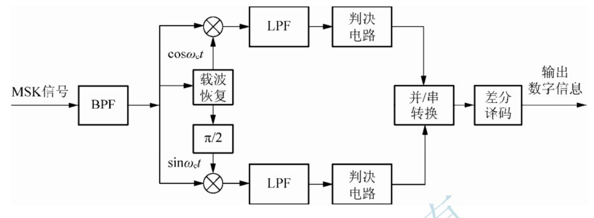


图 4.40　MSK 信号相干解调框图

4. MSK 信号的频谱特性

MSK 信号的归一化双边功率频谱密度的表达式为

$$P_s(f)=\frac{16T_s}{\pi^2}\left[\frac{\cos 2\pi(f-f_c)T_s}{1-16(f-f_c)^2T_s^2}\right]^2 \tag{4-35}$$

其中 f_c 为载频，T_s 为码元宽度。按照上式可以画出 MSK 信号的功率谱曲线如图 4.41 所示。图中还画出了 2PSK 信号的功率谱。由图可以看出，与 2PSK 相比，MSK 信号的功率谱更加紧凑，其第一个零点出现在 $0.75/T_s$ 处，而 2PSK 的第一个零点出现在 $1/T_s$ 处。这说明，MSK 信号功率谱的主瓣所占的频带宽度比 2PSK 信号的窄，当 $(f-f_c)\to\infty$ 时，MSK 的功率谱以 $(f-f_c)^{-4}$ 的速率衰减，它要比 2PSK 的衰减速率大得多。

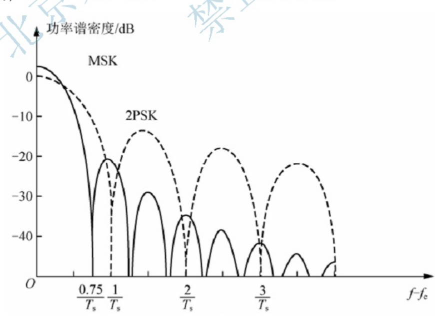
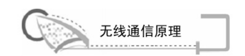


图 4.41　MSK 信号的功率谱



4.3.3 高斯最小移频键控

MSK 信号虽然是包络恒定，带外功率谱密度衰减较快，但 MSK 仍不能满足移动通信环境下对邻道干扰的严格要求。高斯最小移频键控(GMSK)调制方式就是针对该问题而提出的，GMSK 是对 MSK 调制方式的改进，它在 MSK 调制器之前加入一个基带信号预处理滤波器，即高斯低通滤波器，如图 4.42 所示。基带的高斯低通滤波平滑了 MSK 信号的相位曲线，稳定了信号的频率变化，将基带信号变换成高斯脉冲信号，滤除高频分量，其包络无陡峭边沿和拐点，使得发射频谱上的旁瓣水平大大降低，功率谱更加紧凑，大大改善了 MSK 信号的频谱特性，并很好地抑制了移动通信环境中的邻道干扰。如欧洲的 GSM 数字蜂窝移动通信系统采用的就是 GMSK 调制方式。

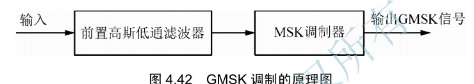


图 4.42 GMSK 调制的原理图

实现 GMSK 信号的调制，关键是设计一个性能良好的高斯低通滤波器，该滤波器必须具有的特性：①有良好的窄带和尖锐的截止特性；②脉冲响应过冲量应尽量小；③输出脉冲响应曲线的面积对应的相位为 $\pi/2$。这些要求主要就是为了抑制高频分量、防止过量的瞬时频率偏移以及满足相干检测的需要。

满足上述特性的高斯低通滤波器，其单位冲激响应为

$$h(t)=\frac{\sqrt{\pi}}{\alpha}\exp\left[-(\alpha\pi t)^{2}\right] \tag{4-36}$$

其中 α 是与高斯滤波器的 3dB 带宽 B_{b} 有关的参数，它们之间的关系为

$$\alpha=\sqrt{\frac{2}{\ln 2}}B_{\mathrm{b}} \tag{4-37}$$

若数字基带信号为双极性不归零的矩形脉冲序列

$$s(t)=\sum_{n}a_{n}b\left(t-nT_{\mathrm{s}}\right),\quad a_{n}=\pm 1 \tag{4-38}$$

其中

$$b(t)=\begin{cases}\dfrac{1}{T_{\mathrm{s}}}，\ |t|\leqslant\dfrac{T_{\mathrm{s}}}{2}\\ 0，其他\end{cases} \tag{4-39}$$

高斯低通滤波器的输出为

$$x(t)=s(t)\times h(t)=\sum_{n}a_{n}g\left(t-nT_{\mathrm{s}}\right),\quad a_{n}=\pm 1 \tag{4-40}$$

其中 $g(t)$ 为高斯低通滤波器的脉冲响应

$$g(t)=b(t)\times h(t)=\frac{1}{T_{\mathrm{s}}}\int_{-T_{\mathrm{s}}/2}^{T_{\mathrm{s}}/2}h(\tau)\mathrm{d}\tau=\frac{1}{T_{\mathrm{s}}}\int_{-T_{\mathrm{s}}/2}^{T_{\mathrm{s}}/2}\frac{\sqrt{\pi}}{\alpha}\exp-\left[\left(\frac{\pi}{\alpha}\tau\right)^{2}\right]\mathrm{d}\tau \tag{4-41}$$

GMSK 信号的表达式为

$$s_{\text{GMSK}}(t)=\cos\left\{\omega_{\text{c}}t+\frac{\pi}{2T_{\text{s}}}\int_{-\infty}^{t}\left[\sum_{n}a_{n}g\left(\tau-nT_{\text{s}}-\frac{T_{\text{s}}}{2}\right)\right]\text{d}\tau\right\} \tag{4-42}$$

高斯低通滤波器的输出脉冲经 MSK 调制得到 GMSK 信号，其相位轨迹由脉冲的形状决定。高斯滤波后的脉冲无陡峭沿，也无拐点，因此 GMSK 信号在码元转换时刻其信号和相位不仅是连续的，而且是平滑的，这样使得 GMSK 信号比 MSK 信号具有更优的频谱特性。GMSK 的解调原理同 MSK 的一样，MSK 信号相干解调器原理图完全适用 GMSK 信号的相干解调。

图 4.43 描述出了 GMSK 信号的功率谱密度。图中，横坐标的归一化频率为$(f-f_{\text{c}})T_{\text{s}}$，纵坐标为谱密度，参变量$B_{\text{b}}T_{\text{s}}$为高斯低通滤波器的归一化 3dB 带宽$B_{\text{b}}$与码元长度$T_{\text{s}}$的乘积。$B_{\text{b}}T_{\text{s}}\to\infty$的曲线是 MSK 信号的功率谱密度，由图可见，GMSK 信号的频谱随着$B_{\text{b}}T_{\text{s}}$值的减小变得紧凑起来。然而，当$B_{\text{b}}T_{\text{s}}$值较小时会引入严重的码间串扰，并且$B_{\text{b}}T_{\text{s}}$值越小，码间串扰越大，GMSK 信号误比特率性能也会变得越差。在实际应用中$B_{\text{b}}T_{\text{s}}$应该折中选择。一般地，当$B_{\text{b}}T_{\text{s}}=0.25$时，误比特率性能下降并不严重。表 4-2 给出了作为$B_{\text{b}}T_{\text{s}}$函数的 GMSK 信号中包含给定功率百分比的带宽。

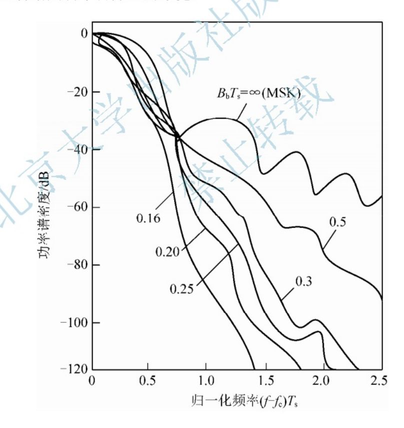


图 4.43　GMSK 信号的功率谱密度

表 4-2　GMSK 包含给定功率百分比的归一化带宽

$B_b T_s$ \ %	90%	99%	99.8%	99.9%
0.2GMSK	0.52	0.79	0.99	1.22
0.25GMSK	0.57	0.86	1.09	1.37
0.5GMSK	0.69	1.04	1.33	2.08
MSK	0.78	1.20	2.76	6.00

GMSK 的解调原理同 MSK 的一样，MSK 信号相干解调器原理图完全适用 GMSK 信号的相干解调。

【例 4.2】为了产生 $B_b T_s = 0.25$ 的 GMSK 信号，当信道数据速率 $R_b = 270\text{kbps}$ 时，试求高斯低通滤波器的 3dB 带宽，并确定射频信道中 90%的功率集中在多大的带宽中。

解：由题目可得 $T_s = \dfrac{1}{R_b} = \dfrac{1}{270\times10^3} \approx 3.7\mu\text{s}$

由已知 $B_b T_s = 0.25$，可知高斯低通滤波器的 3dB 带宽

$$B_b = \frac{0.25}{T_s} = \frac{0.25}{3.7\times10^{-6}} \approx 67.568\ \text{kHz}$$

查表 4-2 可知 90%的功率集中的带宽为

$$B_{90} = 0.57R_b = 0.57\times270\times10^3 = 153.9\text{kHz}$$

4.3.4　四相绝对移相调制

四相绝对移相调制(QPSK)是利用载波的 4 种不同相位来表征数字信息。由于每一种载波相位代表两个比特信息，故每个四进制码元又被称为双比特码元。在 QPSK 实现过程中，首先对输入数据进行串/并转换，将二进制数据每两个比特分为一组，一共有 4 种组合，通常是按格雷码排列即 00、01、11 和 10。然后每组前一比特为同向分量 I，后一比特为正交分量 Q，利用同向分量、正交分量分别对两个正交的载波进行 2PSK 调制，最后将结果叠加就得到 QPSK 信号。

QPSK 信号的表达式为

$$S_{\text{QPSK}}(t) = I(t)\cos\omega_c t - Q(t)\sin\omega_c t \tag{4-43}$$

双比特码元与载波相位的关系见表 4-3，QPSK 信号的相位星座图如图 4.44 所示，其相位每隔 $2T_s$ 跳变一次。图 4.44(a)表示 A 方式时 QPSK 信号的相位星座图，图 4.44(b)表示 B 方式时 QPSK 信号的相位星座图。由图 4.44 可知，QPSK 信号的相位在 4 种可能的相位上跳变。当跳变量为±π时发生相位翻转，引起最大包络起伏。

表 4-3　双比特码元与载波相位关系

双比特码元		载波相位	
a	b	A 方式	B 方式
0	0	225°	0°
1	0	315°	90°
1	1	45°	180°
0	1	135°	270°

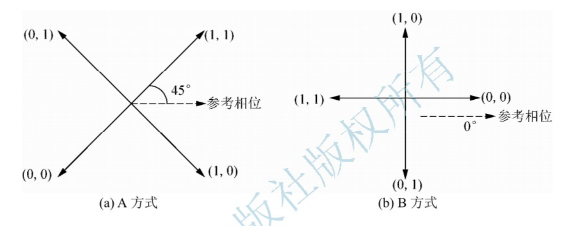


图 4.44　QPSK 信号的相位星座图

典型的 QPSK 信号生成原理图如图 4.45 所示，首先将二进制信息比特流分成两路(同相分量和正交分量)，每一路的比特率 $R_s = R_b / 2$，两路信号再分别与两个正交载波进行相乘后求和产生 QPSK 信号。由于 QPSK 可以看成是两个正交 2PSK 信号的合成，故它可以采用与 2PSK 信号类似的解调方法进行解调，QPSK 相干解调原理框图如图 4.46 所示。

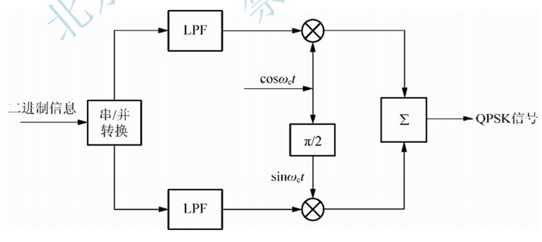


图 4.45　QPSK 信号生成原理图

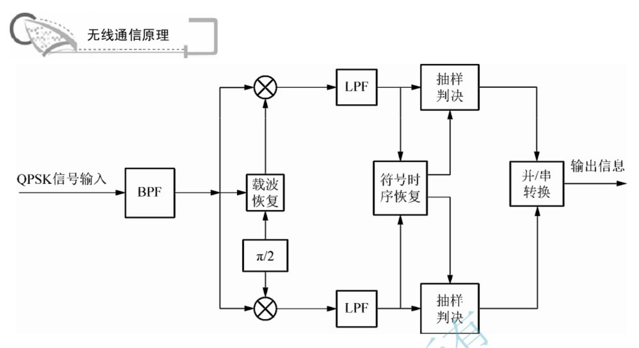


图 4.46　QPSK 相干解调原理图

4.3.5　偏移四相移相键控

QPSK 调制存在诸多问题，如：调制信号带宽为无穷宽，而实际的信道带宽总是有限的；码组中两个比特同时变化时有相位翻转现象，引起包络起伏；包络起伏会导致频谱扩散，增加邻道干扰。为了克服 QPSK 调制存在的上述问题，提出了偏移四相移相键控(Offset-QPSK，OQPSK)调制方式。OQPSK 是 QPSK 的改进型，它与 QPSK 有同样的相位关系，也是把输入码流分成两路，然后进行正交调制。OQPSK 与 QPSK 的不同之处在于它将同相和正交两支路的码流在时间上错开了半个码元周期。由于两支路码元半周期的偏移，每次只有一路可能发生极性翻转，不会发生两支路码元极性同时翻转的现象。因此，OQPSK 信号相位只能跳变 0°、±90°，不会出现 180°的相位跳变。

OQPSK 调制表达式

$$s_{\mathrm{OQPSK}}(t) = I(t)\cos\omega_c t - Q\left(t - T_s/2\right)\sin\omega_c t \tag{4-44}$$

其中 $I(t)$ 表示同相分量；$Q\left(t-T_s/2\right)$ 表示正交分量，它相对于同相分量偏移 $T_s/2$。OQPSK 调制原理图如图 4.47 所示。

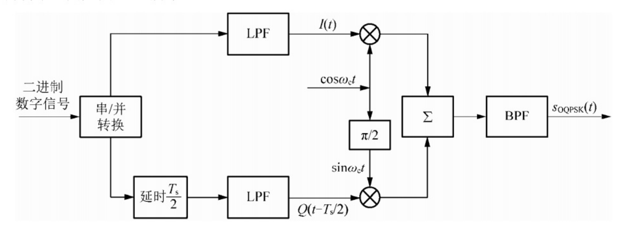


图 4.47　OQPSK 调制原理图



由于同相分量和正交分量不能同时发生变化，相邻一个比特信号的相位只可能发生 ±π/2 的变化，从而消除了相位翻转 ±π 的现象。OQPSK 的 I 、Q 信道波形及相位路径如图 4.48 所示。

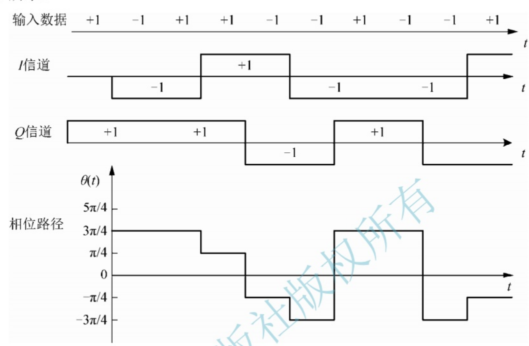


图 4.48　OQPSK 的 I 、Q 信道波形及相位路径

从图 4.49 星座图和相位转移图中看出，OQPSK 调制方式中 ±π 相位的跳变消除了，所以 OQPSK 信号的带限不会导致信号包络经过零点。由于 OQPSK 包络的变化减小了很多，因此对 OQPSK 的硬限幅或非线性放大不会再生出严重的频带扩展，也就是说 OQPSK 即使在非线性放大后仍能保持其带限的性质，因而 OQPSK 非常适合移动通信系统。

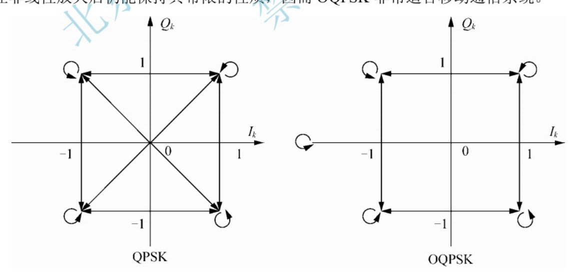


图 4.49　QPSK 和 OQPSK 的星座图和相位转移图

OQPSK 同样可以采用相干方式进行解调，其解调原理与 QPSK 基本相同。比较 OQPSK 和 QPSK，两种调制方式均采用相干解调，理论上误码性能相同。频带受限的 OQPSK 信

号包络比频带受限的 QPSK 信号的小，经限幅放大后功率谱展宽的少，所以 OQPSK 的性能优于 QPSK。实际中，OQPSK 比 QPSK 应用更广泛。然而，OQPSK 信号不能接受差分检测，接收机的设计比较复杂。

4.3.6 π/4-DQPSK 调制

π/4-DQPSK 调制是一种正交移相键控调制技术，它是采用差分编码的π/4-QPSK，并综合了 QPSK 和 OQPSK 两种调制方式的优点。π/4-DQPSK 将 QPSK 的最大相位跳变±π 降为±3π/4，并且比 QPSK 有更好的恒包络特性，在多径扩展和衰落情况下，π/4-DQPSK 的性能比 OQPSK 更好。此外，π/4-DQPSK 既可以采用相干解调也可以采用非相干解调，从而大大简化了接收机的设计。目前，π/4-DQPSK 已用于美国的 IS-136 数字蜂窝系统和个人接入通信系统(PACS)、日本的个人数字蜂窝系统(PDC)。

π/4-DQPSK 可采用正交调制方式产生，其原理如图 4.50 所示。输入数据首先经串/并转换后得到两路序列 I_k 和 Q_k，它们的码元速率等于输入串行比特速率的一半；然后这两路数据通过差分相位编码器进行相位差分编码，在 $(k-1)T_s \leqslant t \leqslant kT_s$ 期间内，输出信号 U_k 和 V_k；为了抑制已调信号的旁瓣，在输出信号与载波相乘之前，通常还经过具有升余弦特性的成形滤波器(LPF)；最后再分别和一对正交载波相乘进行正交调制合成，就得到了π/4-DQPSK 信号。

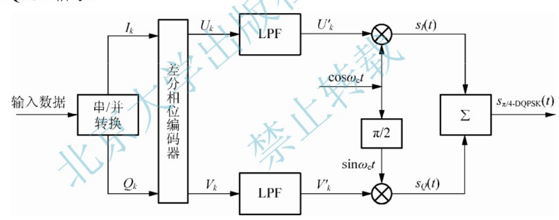


图 4.50　π/4-DQPSK 调制原理图

由于π/4-DQPSK 信号的相位跳变取决于相位差分编码，下面对相位差分编码与π/4-DQPSK 信号相位跳变的关系进行分析。设π/4-DQPSK 信号表达式为

$$s_{\pi/4\text{-DQPSK}}(t)=\cos\left(\omega_c t+\varphi_k\right) \tag{4-45}$$

将上式展开，可得

$$s_{\pi/4\text{-DQPSK}}(t)=\cos\omega_c t\cos\varphi_k-\sin\omega_c t\sin\varphi_k=U_k\cos\omega_c t-V_k\sin\omega_c t \tag{4-46}$$

其中 φ_k 是当前码元的附加相位，它是前一码元的附加相位 φ_{k-1} 与当前码元相位增量 $\Delta\varphi_k$ 之和，即：$\varphi_k=\varphi_{k-1}+\Delta\varphi_k$，则有

$$\begin{cases}U_k=\cos\varphi_k=\cos\left(\varphi_{k-1}+\Delta\varphi_k\right)=\cos\varphi_{k-1}\cos\Delta\varphi_k-\sin\varphi_{k-1}\sin\Delta\varphi_k\\V_k=\sin\varphi_k=\sin\left(\varphi_{k-1}+\Delta\varphi_k\right)=\sin\varphi_{k-1}\cos\Delta\varphi_k+\cos\varphi_{k-1}\sin\Delta\varphi_k\end{cases} \tag{4-47}$$



其中$U_{k-1}=\cos\varphi_{k-1}$，$V_{k-1}=\sin\varphi_{k-1}$，则上式可写为

$$\begin{cases}U_k=U_{k-1}\cos\Delta\varphi_k-V_{k-1}\sin\Delta\varphi_k\\V_k=V_{k-1}\cos\Delta\varphi_k+U_{k-1}\sin\Delta\varphi_k\end{cases}\tag{4-48}$$

式(4-48)是π/4-DQPSK 的一个基本关系式，它表明了前一码元两正交信号U_{k-1}、V_{k-1}与当前码元两正交信号U_k和V_k之间的关系，它取决于当前码元的相位增量$\Delta\varphi_k$，而$\Delta\varphi_k$又取决于相位编码器的输入码组I_k和Q_k，相位差分编码就是将I_k和Q_k的 4 个状态用 4 个$\Delta\varphi_k$值来表示，其相位跳变规律见表 4-4。

表 4-4　π/4-DQPSK 的相位跳变规率

I_k	Q_k	$\Delta\varphi_k$	$\cos\Delta\varphi_k$	$\sin\Delta\varphi_k$
1	1	π/4	$1/\sqrt{2}$	$1/\sqrt{2}$
0	1	3π/4	$-1/\sqrt{2}$	$1/\sqrt{2}$
0	0	−3π/4	$-1/\sqrt{2}$	$1/\sqrt{2}$
1	0	−π/4	$1/\sqrt{2}$	$1/\sqrt{2}$

由表 4-4 可看出，在码元转换时刻$\Delta\varphi_k$只有±π/4 和±3π/4 这 4 种取值，所以相位φ_k有 8 种可能的取值，π/4 -DQPSK 的相位关系如图 4.51 所示。从图中可以看出其星座图的 8 个点实际是由两个彼此偏移π/4 的 QPSK 星座图构成的，相位的跳变总是在这两个星座图之间交替进行，信号相位跳变必定在图中的“。”组和“×”组之间进行，即在相邻码元仅会出现从“。”组到“×”组相位点(或“×”组到“。”组)的跳变，而不会在同组内跳变，并且所有的相位路径都不经过原点(圆心)。这种特性使得π/4 -DQPSK 信号的包络波动比 QPSK 的要小，进而降低了最大功率和平均功率的比值。同时也可以看到，U_k和V_k可能的取值只有 0、$\pm1/\sqrt{2}$、±1 这 5 种，它们分别对应于图 4.51 中 8 个相位点的坐标值。

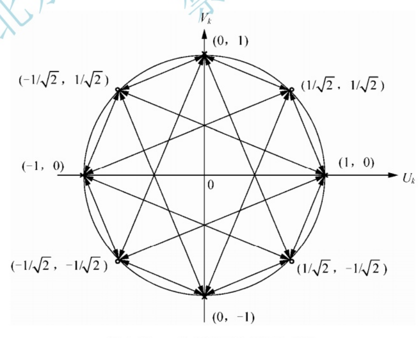


图 4.51　π/4 -DQPSK 的相位关系

π/4-DQPSK 信号解调可采用相干解调和非相干解调的方法。为了便于实现，经常采用差分检测的方法来解调 DQPSK 信号。在低比特率、快速瑞利衰落信道中，由于不依赖相位同步，差分检测提供了较好的误码性能。π/4-DQPSK 信号基带差分检测器的原理图如图 4.52 所示。在解调器中，本地正交载波要求与信号的发射载波频率相同，并不要求相位完全相同，可以允许有一定的相位差，这个相位差可以在差分检测过程中消去。同相和正交两个支路分别与本地载波相乘后经低通滤波，再进行差分译码，译码后的信号根据调制时的相位跳变规则进行判决，将判决结果再经并/串转换即可恢复所传输的数据。

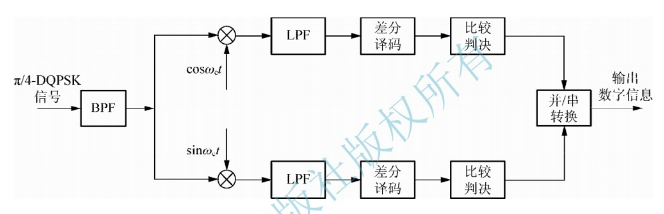


图 4.52　基带差分解调原理图

π/4-DQPSK 信号的解调还可以采用中频差分检测，其原理框图如图 4.53 所示。中频差分检测时，信号同样分成两路：一路是信号本身和信号延迟一个码元后的信号两者相乘，另一路是信号经移相 π/2 后与其延迟一个码元后的信号两者相乘，经两个支路分别相乘后的信号再经低通滤波，得到上、下两支路的低频分量，后面的判决过程则与基带差分检测完全一样。此方法的优点是不需要本地产生载波。

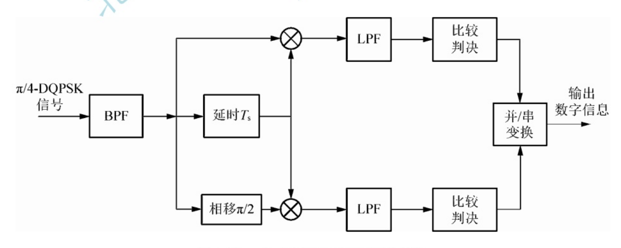


图 4.53　π/4-DQPSK 的中频差分解调



4.3.7　正交频分复用调制

1. OFDM 基本原理

前面介绍的 ASK、FSK、PSK、QAM、MSK 等调制方式，在某一时刻都只用单一的载波频率来传输数字信息。而实际中，为了提高信道利用率，可以采用多载波调制方式来实现信息的有效传输。传统的多载波调制系统是将待发送的高速率信息数据流经串/并转换，分割为若干路低速率并行数据流，然后每路低速率数据采用一个独立的载波调制，并采用频分复用调制方式叠加在一起构成发送信号。传统多载波调制子载波的频率设置如图 4.54(a)所示，这种方式只是将一个信道的带宽分成多个窄带子信道，每个信道采用单一载波，各个载波所传输的数字信息都来自于同一信源。接收端首先要采用滤波器组进行各子带的分离，然后解调各路低速率数据，再经并/串转换，恢复出高速率的信息数据流。这种方案的优点是实现简单、直接；缺点是频率利用率不高，各子载波间要保留保护带，且多滤波器的实现也较困难。

正交频分复用(Orthogonal Frequency Division Multiplexing，OFDM)是一种特殊的多载波调制技术，它的多载波之间有一定的约束要求，即各子载波频谱有 1/2 重叠且保持相互正交，如图 4.54(b)所示。OFDM 中各子信道在时间上相互正交，在频率上相互重叠，接收端通过相干解调技术分离出各子载波，同时又消除了码间干扰的影响，这样就可以节省带宽而获得更高的频带效率，并且避免使用滤波器组。

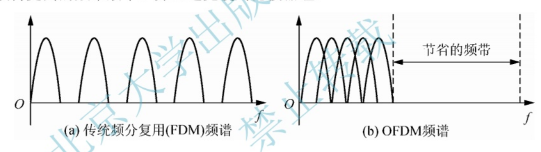


图 4.54　FDM 与 OFDM 信道划分方式对比

OFDM 调制的基本原理是将串行的高速信息数据流转换成若干并行的低速数据流，并行地调制在多个相互正交的子载波上同时进行传输，然后所有子载波叠加在一起构成发送信号。OFDM 调制可以降低每个子载波的码元速率，增大码元的码元周期，使系统的抗衰落和抗干扰能力大大提高，且频带利用率更高，所以 OFDM 非常适合移动场合中的数据高速传输。

OFDM 调制系统原理如图 4.55 所示，待发送的高速串行数据流经串/并转换后，得到 N 路并行低速数据，这 N 路并行码分别调制在 N 个相互正交的子载波上，相加后得到 OFDM 信号

$$s_{\text{OFDM}}(t)=\sum_{k=0}^{N-1}d(k)\cos\left(2\pi f_k t\right),\quad 0<t<T_{\text{s}} \tag{4-49}$$

其中 $d(k)$ 为第 k 路并行码，f_k 为第 k 路子载波频率，并且有 $f_k=f_0+k\Delta f$，$\Delta f=\dfrac{1}{T_{\text{s}}}$ 为子载波的频率间隔，它是 OFDM 调制系统的重要设计参数。

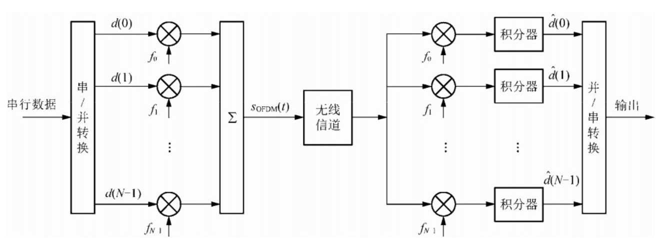


图 4.55　OFDM 调制系统原理框图

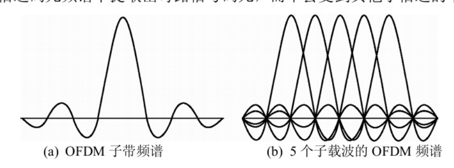
为了保证各子载波间两两相互正交，则有

$$\frac{1}{T_s}\int_0^{T_s}\cos(2\pi f_i t)\cos(2\pi f_j t)\mathrm{d}t=\begin{cases}1, & i=j\\0, & i\neq j\end{cases} \tag{4-50}$$

即在信道传输码元的持续时间内，不同子载波乘积的积分值为零。由于子载波的相互正交性，使得接收端在理论上很容易实现对某个子载波信号的相干解调，而不受其他子载波的干扰。接收到的信号同时进入 N 个并行支路，分别与 N 路子载波相乘和积分(相干解调)，便可以恢复并行支路的数据

$$\hat{d}(k)=\int_0^{T_s}s_{\mathrm{OFDM}}(t)\cos(2\pi f_k t)\mathrm{d}t=\int_0^{T_s}\sum_{k=0}^{N-1}d(k)\left[\cos(2\pi f_k t)\right]^2\mathrm{d}t=d(k) \tag{4-51}$$

OFDM 信号由 N 路信号叠加而成，当 OFDM 信号各子路的码型选择单极性不归零矩形脉冲时，每路信号的频谱都是以子载波频率为中心频率的 Sa 函数形式，且相邻信号频谱之间有 $1/T_s$ 宽度的重叠。OFDM 信号的频谱结构如图 4.56 所示，图中，在每个子载波频率处，该子载波频谱取最大值，而所有其他子载波的频谱值恰好此时均为零，因此可以从多个相互重叠的子信道码元频谱中提取出每路信号码元，而不会受到其他子信道的干扰。



(a) OFDM 子带频谱　　(b) 5 个子载波的 OFDM 频谱

图 4.56　OFDM 信号的频谱结构

忽略旁瓣功率，OFDM 信号的频谱宽度为

$$B_{\mathrm{OFDM}}=(N-1)\frac{1}{T_s}+\frac{2}{T_s}=\frac{N+1}{T_s} \tag{4-52}$$

OFDM 信道在 T_s 时间内传输 N 个并行码元，则码元速率为 $R_B = N/T_s$，若每个支路采用 M 进制调制，则比特速率为 $R_b = R_B \log_2 M$，对应频带利用率为

$$\eta_{OFDM} = \frac{R_b}{B_{OFDM}} = \frac{N}{N+1}\log_2 M \tag{4-53}$$

2. OFDM 的数字化实现

模拟实现的 OFDM 虽然已经可以通过频谱混叠实现频带利用率的提升，但是所需要的设备非常复杂，特别是当 N 很大时，每个子载波都需要有单独的调制器和解调器，需要大量的乘法器等设备，宽带信道提供数百上千个子带，时域实现较为困难且费用非常昂贵。现代数字信号处理器的广泛应用使 OFDM 调制解调的实现可以采用 IFFT 和 FFT 变换，将多载波转化为单载波系统，从而大大简化了系统实现的复杂度。基于 IFFT/FFT 的 OFDM 系统原理图如图 4.57 所示，OFDM 信号的产生是基于快速傅里叶逆变换实现发送信号的处理，而接收端用快速傅里叶正变换处理信号。

在发送端，调制信息经过串/并转换，转换为 N 路并行的比特流，对应于 N 个不同的子载波，每个子载波载荷一个码元信息。信号映射是根据某种调制模式进行星座点的映射，生成 N 个调制码元，将一个码元的多比特信息映射为某种调制模式的星座图上的某一星点，该星点的信息即是该码元对应的子载波上的调制码元。信号映射后经 IFFT 模块完成多载波调制。为了在接收端有效地抑制码间干扰(ISI)，通常在每一个 OFDM 码元前要附加上长度为 N_g 个采样的保护间隔，一般使用循环前缀作为保护间隔。在 OFDM 码元之间插入保护间隔，令保护间隔的长度大于无线信道的最大时延扩展，可以最大限度地消除多径带来的码间干扰，又可以避免由多径带来的子信道间干扰。最后经并/串转换后送入 D/A 转换形成双极性多电平方波，再经上变频调制最后形成 OFDM 信号发送出去。

接收端的主要任务是完成发送部分的相反过程。首先经下变频后，通过 A/D 转换将从信道中获得的模拟信号转换为数字信号，再经串/并转换、去除循环前缀，将得到的数据进行 FFT 变换解调，获得各子载波上的接收信号，然后通过星座点逆映射得到每个子载波上的接收比特，最后再经过并行到串行的转换恢复出原始比特流。

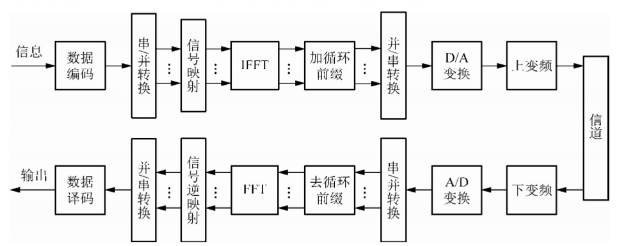


图 4.57　基于 FFT/IFFT 的 OFDM 系统原理图

3. OFDM 的特点

OFDM 调制是一种高效调制技术。相比一般的数字通信系统，OFDM 调制具有较强的抗多径传播和抗频率选择性衰落的能力以及较高的频带利用率等优点，因而适合于高速率的无线通信系统。OFDM 现已成功地应用于接入网中的高速数字环路 HDSL、非对称数字环路 ADSL、无线本地环路、无线本地局域网、高清晰度电视 HDTV 的地面广播系统。OFDM 也是第三代、第四代移动通信系统中采用的核心技术。

OFDM 技术的优点如下。

(1) 可有效解决多径问题，适用于多径环境和衰落信道中的高速数据传输。

(2) 基于 FFT/IFFT 的 OFDM 系统实现难度远低于均衡器的单载波系统。

(3) 可以很好地抵抗窄带干扰，有较强的抗频率选择性衰落能力。

(4) 信道利用率很高，这一点在频谱资源有限的无线环境中尤为重要。

(5) 在相对低时变信号中，根据特定子载波的信噪比来调整每个子载波的数据传输速率，可以有效地增加容量。

(6) 支持动态比特分配及动态子信道分配方法。

(7) 融合能力强，易于与其他多址接入方式相结合使用。

OFDM 的不足之处如下。

(1) OFDM 对相位噪声和载波频偏十分敏感。

(2) 多载波信号是多个单载波信号的叠加，因此其峰值功率与平均功率的比值较大，对非线性放大更为敏感，对前端放大器的线性要求较高。

(3) OFDM 系统对同步系统的精度要求更高，大的同步误差不仅造成输出信噪比的下降，还会破坏子载波间的正交性，造成载波间干扰，从而大大影响系统的性能，甚至使系统无法正常工作。

4.4 扩频调制技术

4.4.1 扩频调制基本原理

扩频通信的一般原理如图 4.58 所示。在发送端输入的信息经信息调制形成数字信号，然后由扩频码发生器产生的扩频码序列去调制数字信号以展宽信号的频谱，展宽以后的信号再对载频进行调制(如 PSK 或 QPSK、OQPSK 等)，最后通过射频功率放大送到天线上发射出去。在接收端，从接收天线上收到的宽带射频信号，经过输入电路、高频放大器后送入变频器，下变频至中频，然后由本地产生的与发端完全相同的扩频码序列去解扩，最后经解调，恢复原始信息。一般的扩频通信系统都要进行 3 次调制和相应的解调。一次调制为信息调制，二次调制为扩频调制，三次调制为射频调制，以及相对应的信息解调、解扩和射频解调。

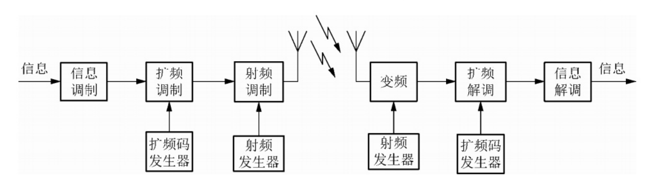


图 4.58　扩频通信原理框图

扩频调制技术是扩频通信中的关键技术，扩频调制和解扩是扩频通信的核心部分，也是扩频通信系统与普通数字通信系统主要的区别所在。扩频调制的基本原理是将发送的信息扩展到一个很宽的频带上，通常使用的传输带宽比要求的最小信号带宽大几个数量级，频带的扩展由独立于信息的扩频码实现；在接收端则用相同的扩频码进行相关解调，实现解扩和恢复所发送的信息。尽管扩频通信系统对单个用户频带利用率很低，但是扩频调制的优点是很多用户可以同时使用同一个带宽，而不会产生明显的相互干扰。在无线移动环境的应用中，扩频调制系统具有许多优良的特性，如系统的抗干扰性能非常好，抗衰落和抗阻塞的能力较强，具有选择地址(用户)的能力，信号的功率谱密度较低，信号隐蔽性较好并且功率污染较小，比较容易进行数字加密，防止窃听，在共用信道中能实现码分多址复用，多用户共享相同的信道，无须进行频率规划。扩频通信还具有良好的抗多径干扰的能力，使无线通信的性能变得更加可靠。扩频通信主要应用于军事抗干扰通信、数字蜂窝移动通信系统、卫星移动通信、室内无线通信和未来个人通信等领域，还可用于定时、定位及测距，应用在导航、雷达等系统中。

扩频调制系统按其工作方式可分为：直接序列(DS)扩频、跳频(FH)扩频、跳时(TH)扩频以及上述几种方式的组合扩频。

4.4.2　直接序列扩频

直接序列扩频(Direct Sequency Spread Spectrum，DS-SS)是指直接用具有高码率的扩频码序列在发送端扩展基带信号的频谱。而在接收端，用相同的扩频码序列进行解扩，把展宽的扩频信号还原成原始的信息。直接序列扩频调制和接收机解调原理图分别如图 4.59 和图 4.60 所示。

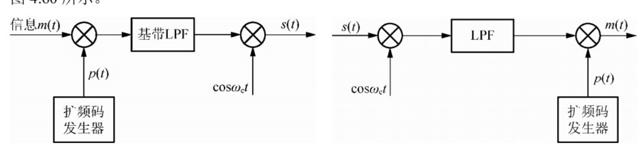


图 4.59　直接序列扩频通信调制原理图　　　　图 4.60　接收机解调原理图

在 DS-SS 系统中，常用调制方式有 2PSK、QPSK 和 MSK 等。最简单的形式就是使用 2PSK 作扩频调制，扩频后的信号在数学上可以认为是 2PSK 信号。

设数字基带信号 $m(t)=\sum_{n=-\infty}^{\infty}a_n g(t-nT_s)$，$a_n$ 为信息码，取值为−1 或+1，T_s 为码元间隔，扩频序列为 $p(t)$，基带信号扩展后的序列为 $d(t)$，经过 2PSK 调制后，输出的扩频信号为

$$s(t)=d(t)\cos\omega_c t=m(t)p(t)\cos\omega_c t \tag{4-54}$$

DS-SS 扩频接收机的主要功能是与接收到的扩频信号保持同步、对信号进行解扩和检测数据。直接序列扩频的解调原理如图 4.60 所示，图中，首先将输入的直扩信号进行 2PSK 解调，然后再与伪随机序列相乘进行解扩。为了正确恢复信号，在接收端产生的扩频序列必须与接收的扩频信号中的扩频序列同步。设解调器输入为

$$r(t)=s(t)+n(t)=m(t)p(t)\cos\omega_c t+n(t) \tag{4-55}$$

其中 $n(t)=n_c(t)\cos\omega_c t-n_s(t)\sin\omega_c t$ 为加性高斯白噪声。

2PSK 解调输出

$$s'(t)=\frac{1}{2}m(t)p(t)+\frac{1}{2}n_c(t) \tag{4-56}$$

经解扩后输出

$$m_o(t)=\frac{1}{2}P_p m(t)+n_o(t) \tag{4-57}$$

其中 P_p 为扩频序列 $p(t)$ 的平均功率。

图 4.61 给出了直接序列扩频对应信号的功率谱密度。若接收端解调器的输入信号是经扩频后宽带有用 2PSK 信号和窄带干扰信号的叠加，如图 4.61(a)所示，与本地扩频码相乘解扩后，有用信号则变为窄带的 2PSK 信号，其功率谱密度受到压缩而使其频谱密度增大，窄带干扰信号的功率谱解扩后带宽展宽，功率谱密度下降，如图 4.61(b)所示。因此，直接序列扩频系统通过接收端进行解扩处理，使原来掩埋在噪声和干扰下的有用信号的功率谱得以增强，经滤波处理后，解调器将大部分信号带外的干扰滤除，使噪声和干扰得到相对抑制，从而提高了扩频系统的抗干扰能力。

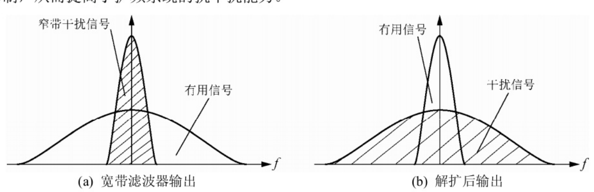


图 4.61 2PSK 直接序列扩频对应信号的功率谱密度

4.4.3 跳频扩频

跳频扩频(Frequency Hopping Spread Spectrum，FH-SS)是用扩频码序列去进行移频键控调制，使载波频率不断地跳变。简单的移频键控如 2FSK 只有两个频率，而跳频系统的



载波频率则有几个、几十个甚至上千个，并由所传信息与扩频码的组合进行选择控制而不断地发生跳变。跳频的模式由扩频码决定，所有可能的载波频率的集合称为跳频集。图 4.62 为跳频扩频的原理示意图。在发送端，信息码序列与扩频码序列经调制后按照不同的码字控制频率合成器，频率合成器的输出频率根据码字的改变而变化，形成了频率的跳变。在接收端，为了解调跳频信号，需要有与发端完全相同的本地扩频码发生器控制本地频率合成器，使其输出的跳频信号能在混频器中与接收信号差频，产生固定的中频信号，然后经中频带通滤波器及信息解调器输出恢复的信息。从跳频扩频原理可以看出，跳频系统也占用了比信息带宽要宽得多的频带，跳频发生的频谱带宽称为总跳频带宽，跳频集合中使用的信道带宽称为瞬时带宽，跳频之间的持续时间称为跳频周期。图 4.63 给出了频率跳变示意图，由图可见，跳频扩频的输出频谱在一定宽频带范围内所选择的某些频率中随机地跳变。

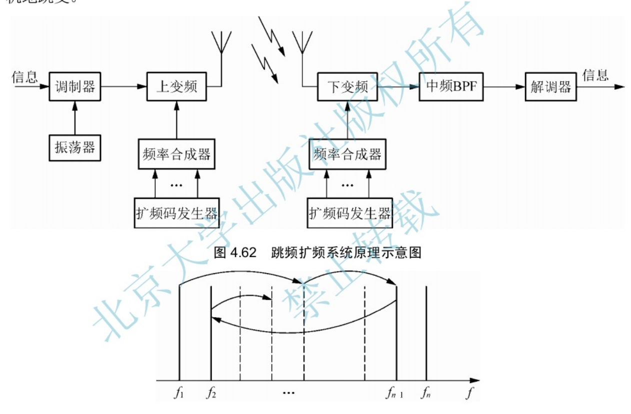
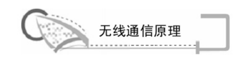


图 4.62 跳频扩频系统原理示意图

图 4.63 跳频扩频系统频率跳变示意图

跳频扩频可以分为快跳扩频和慢跳扩频两种。快跳是指扩频跳频速率大于码元速率，此时，系统在一个码元时间内发生多次频率跳变；慢跳是指频率每跳时间内传送若干码元信息，目前多采用慢跳扩频通信。

跳频扩频通信具有保密性好、抗干扰能力强、系统兼容性好、便于实现多址通信等。由于跳频扩频系统对载波的调制方式没有限制，在军用短波和超短波电台中应用广泛。如在敌我通信对抗中，固定频率的通信很容易暴露目标且易于被截获，而跳频扩频则使用“打一枪换一个地方”的游击策略，不断跳变通信调制的载波频率，使敌方很难判断出通信真正使用的频率，即使被发现了，通信频率也已经很快地换到了另外的频率，从而实现隐蔽、抗干扰和难以被截获的通信目的。



4.4.4 跳时扩频

跳时扩频(Time Hopping Spread Spectrum，TH-SS)是通过时间跳变实现扩展频谱的通信方式，跳时是指使发射信号在时间轴上离散地跳变。在跳时扩频系统中，通常先将时间轴分成帧，每个帧分成许多时隙，数字信号在时隙上使用快速突发脉冲传输，一帧内发射信号的时隙由扩频码序列控制。因此，跳时也可以理解为用一定扩频码序列进行选择的多时隙的时移键控。由于采用了很窄的时隙发送信号，相对来说，信号的频谱也就展宽了。图 4.64 是跳时扩频系统的原理图。图 4.65 所示为 TH-SS 信号时间-频率图，可以看出，TH-SS 是占用整个频段的一小段时间，而不是在全部时间里使用部分频段。

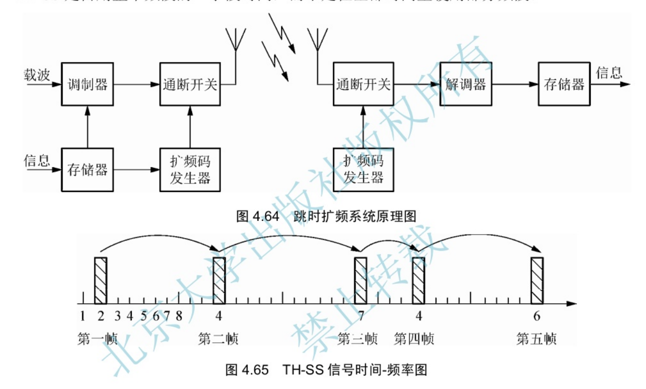


图 4.64　跳时扩频系统原理图

图 4.65　TH-SS 信号时间-频率图

跳时扩频系统由扩频码序列控制时隙的位置按一定规律跳变，用时间的合理分配来减少时分复用系统之间的干扰和避开附近发射机的强干扰，其优点是减少了工作时间的占空比，其缺点是系统对定时要求较高。

以上 3 种扩频方式是最基本的方式，直接序列扩频和跳频扩频是最常用扩频方式，实际中为了发挥每种扩频技术的优势，往往还将这 3 种扩频方式进行组合构成混合扩频方式来使用。

本 章 小 结

本章主要介绍了无线通信调制技术，调制是将基带信号加载到高频载波上的过程，其目的是将需要传输的模拟信号或数字信号变换成适合信道传输的频带信号，以满足无线通信对信息传输的基本要求。本章简要介绍了无线调制技术的基本概念和分类，根据基带信



号、载波和调制器功能的不同，调制技术可以有多种分类。根据基带信号的不同，调制可分为模拟调制和数字调制；根据载波信号的不同，调制可分为连续波调制和脉冲调制；根据基带信号对载波的控制参量的不同，调制可分为幅度调制、频率调制和相位调制。

本章重点介绍了基本的现代数字调制技术 2ASK、2FSK、2PSK 和 2DPSK 调制的基本原理和调制解调过程。为了提高频谱利用率和信息传输速率，无线通信往往采用高性能的数字调制技术：QAM 调制、最小移频键控 MSK 调制、高斯最小移频键控 GMSK 调制、QPSK 调制、OQPSK 调制、DQPSK 调制、OFDM 调制等技术，本章对这几种高性能的数字调制技术的原理应用和特点做了重点介绍。在无线移动通信环境的应用中，扩频调制系统具有许多优良的特性，如系统的抗干扰性能非常好，抗衰落和抗阻塞的能力较强，信号隐蔽性较好并且功率污染较小，比较容易进行数字加密，防止窃听，使无线通信的性能变得更加可靠。扩频通信主要应用于军事抗干扰通信、数字蜂窝移动通信系统、卫星移动通信、室内无线通信和未来个人通信等领域。

习　题

4.1 填空题

(1) 常用的数字基带信号波形有__________、__________、__________、__________、__________和__________等。

(2) 扩频调制系统按其工作方式可分为：__________、__________、__________以及上述几种方式的组合。

(3) 2PSK 调制信号在解调时会产生__________现象。

4.2 什么是调制？调制技术有哪些分类？

4.3 设发送的二进制信息为 1100011101，采用 2ASK 调制方式，已知码元传输速率为 1200 波特，载波频率为 2400Hz，试设计构成一种 2ASK 信号调制器原理框图，并画出对应的 2ASK 信号的时域波形。

4.4 设发送的二进制信息为 1100111001，采用 2FSK 调制方式，已知码元传输速率为 1000 波特，1 码元的载波频率为 2000Hz，0 码元的载波频率为 3000Hz，试设计构成一种 2FSK 信号调制器原理框图，并画出对应的 2FSK 信号的时域波形。

4.5 设发送的二进制信息为 1110100011，采用 2PSK 调制方式，已知码元传输速率为 1000 波特，载波频率为 1000Hz，试设计构成一种 2DPSK 信号调制器原理框图，并画出对应的 2DPSK 信号的时域波形。

4.6 设发送的二进制信息为 1011100101，采用 2DPSK 调制方式，已知码元传输速率为 1000 波特，载波频率为 1000Hz，试设计构成一种 2PSK 信号调制器原理框图，并画出对应的 2PSK 信号的时域波形。

4.7 图 4.66 是两种 8QAM 信号星座图，相邻信号点的最小距离为 d，假设各信号点是等概率的。试分别求两个星座信号的平均功率。

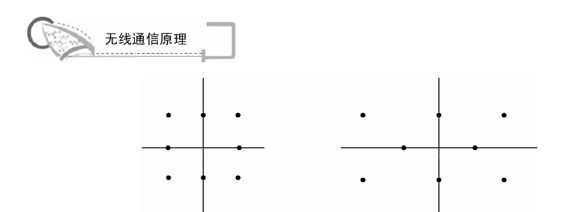


图 4.66　两种 8QAM 信号星座图

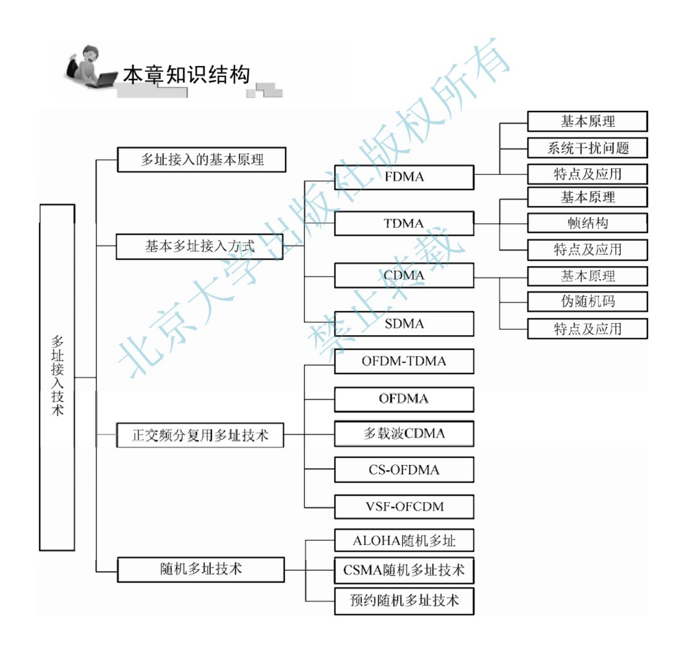
4.8 已知发送数据序列为 1011001011，传输速率为 128Kbps，载波频率为 256kHz，试画出 MSK 信号附加相位路径图；画出 MSK 信号时间波形。

4.9 什么是 OFDM 调制？简述 OFDM 调制的基本原理及其优缺点。

4.10 简述扩频调制的基本原理及其特点。

第 5 章
多址接入技术

本章知识结构



本章教学目的与要求

- 掌握多址接入的基本概念和多址方式。
- 掌握 FDMA 技术的原理及系统的特点，了解 FDMA 系统中的干扰问题。
- 掌握 TDMA 技术的原理及系统的特点，熟悉 TDMA 的帧结构，了解 TDMA 系统的同步与定时。
- 掌握 CDMA 技术的原理及系统的特点，熟悉正交 Walsh 函数、m 序列、Gold 序列
- 了解空分多址(SDMA)技术的原理。
- 掌握系统容量的定义，熟悉 FDMA、TDMA 和 CDMA 系统容量的分析与比较。



引言

无线通信是一个多信道同时工作的系统，具有广播信道和大面积无线电磁波覆盖的特点，如何建立多用户之间的无线连接，以满足不同用户同时进行通信的要求，实现多个用户地址的动态划分、动态寻址和识别，使多用户共享系统资源，这便是多址接入问题，解决多址接入问题的方法就是多址接入技术。多址接入技术是无线通信中的重要技术，在无线通信中占有重要地位，其实质是对所使用的信道或信号进行分割以供不同地址的用户所使用，它关系到小区的构成、系统的容量、频谱和信道利用率以及系统复杂性等诸多问题。本章将主要介绍无线通信系统中的多址接入技术，如频分多址(FDMA)、时分多址(TDMA)、码分多址(CDMA)、空分多址(SDMA)、正交频分多址和随机多址等技术。

【案例 5.1】CDMA 手机

CDMA 手机即采用了码分多址技术(Code Division Multiple Access，CDMA)的手机，最初 CDMA 手机在中国内陆的通信运营商为中国联通，现在为中国电信。与 GSM 手机相比，CDMA 手机可以说是名副其实的绿色手机，它发射功率极小(2mW)，只是我们现在使用的 GSM 手机(功率为 125mW)的 1/60，甚至低于电视屏幕产生的辐射功率。CDMA 手机采用了先进的切换技术——软切换技术，使得 CDMA 手机的通话没有 GSM 手机的掉线现象，可以与固定电话相媲美。使用 CDMA 网络，运营商的投资相对减少，这就为 CDMA 手机资费的下调预留了空间。CDMA 手机采用以扩频通信为基础的调制和多址通信方式，其容量比模拟技术高 10 倍，超过 GSM 网络约 4 倍。CDMA 中所提供的语音编码技术，通话品质比目前的 GSM 好，可将用户对话时周围环境噪音降低，使通话更清晰。就安全性能而言，CDMA 不但有良好的认证体制，更因其传输特性，用地址码来区分用户，防盗听能力大大增强。更为重要的是，基于宽带技术的 CDMA 使得移动通信中视频应用成为可能，从而使手机从只能打电话和发送短信息等狭窄的服务中走向宽带多媒体应用。

【案例 5.2】3G 的技术基础——CDMA

国际电信联盟(ITU)确定 WCDMA、CDMA2000 和 TD-SCDMA 为第三代移动通信(3G)的三大主流技术，并写入 3G 技术指导性文件《2000 年国际移动通信计划》(简称 IMT-2000)。CDMA 码分多址技术是第三代移动通信的技术基础。CDMA 系统以其频率规划简单、系



统容量大、通信质量高、软容量和软切换等特性显示出巨大的发展潜力。WCDMA 即宽带 CDMA 技术，它是基于 GSM 网发展出来的 3G 技术规范，是欧洲提出的宽带 CDMA 技术。CDMA2000 是由窄带 CDMA(CDMA IS95)技术发展而来的宽带 CDMA 技术，也称为 CDMA Multi-Carrier，由美国高通北美公司为主导提出，摩托罗拉、Lucent 和后来加入的韩国三星都有参与，韩国现在成为该标准的主导者。TD-SCDMA 全称为 Time Division-Synchronous CDMA(时分同步 CDMA)，该标准是由中国大陆独自制定的 3G 标准，1999 年 6 月 29 日，中国原邮电部电信科学技术研究院(大唐电信)向 ITU 提出。该标准将智能天线、同步 CDMA 和软件无线电等当今国际领先技术融于其中，在频谱利用率、对业务支持灵活性、频率灵活性及成本等方面具有独特的优势。

5.1　多址技术的原理

5.1.1　多址技术的基本原理

多址技术是无线通信的关键技术之一，多址技术主要研究如何将有限的通信资源在多个用户之间进行有效地切割与分配，在保证多用户之间通信质量的同时尽可能地降低系统的复杂度，并获得较高的系统容量。多址技术中对通信资源的切割和分配也就是对多维无线信号空间的划分，对信号在不同的维上进行划分就对应着不同的多址技术，常见的维有信号的时域、频域和空间域，此外还有信号的各种扩展维。多址技术的数学基础是信号的正交分割原理，多址划分就是区分不同的用户地址，这需要利用射频频段辐射的电磁波来寻找动态的用户地址，赋予各个信号不同的特征，也就是打上不同的“地址”，然后根据各个信号特征之间的差异来区分，按“地址”分发，实现互不干扰的通信。为了实现多址信号之间互不干扰，信号之间必须满足正交特性。

设一组相互正交的信号为

$$x(t)=\sum_{i=1}^{n}\lambda_i x_i(t) \tag{5-1}$$

其中 $x_i(t)$ 为第 i 个用户的信号；λ_i 为第 i 个用户 $x_i(t)$ 的正交参量，且正交参量应满足

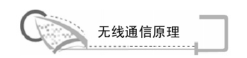
$$\lambda_i\cdot\lambda_j=\begin{cases}1, & \text{当}\,i=j\text{时}\\ 0, & \text{当}\,i\neq j\text{时}\end{cases} \tag{5-2}$$

其中 λ_i 和 λ_j 分别为对应用户 i 和 j 的信号。信号的正交性是通过信号正交参量来实现的，当正交参量仅考虑时间、频率或码型时，信号可以表示为时间、频率和码型的函数，即写为

$$s(c,f,t)=c(t)s(f,t) \tag{5-3}$$

其中 $c(t)$ 是码型函数，$s(f,t)$ 是时间和频率的函数。按照这 3 个参量的分割，可以实现时分多址、频分多址、码分多址技术，在其他各种扩展维上进行信号空间划分的方法在原理上都与此类似。

多址技术的关键是设计具有正交性的信号集合，使各信号相互无关。在实际工作中，要做到完全正交和不相关是比较困难的，一般采用准正交，即互相关很小，允许各信号之间存在一定干扰，但要将干扰控制在允许范围内。在无线通信中，多个不同地址的用户之



间建立通信链路必须采用某种多址接入技术，如在移动通信中，许多移动台要同时通过一个基站和其他移动台进行通信，必须区分出是哪一个移动台发出来的信号，那么基站是以怎样的信号传输方式接收、处理和转发移动台来的信号呢？基站又以怎样的信号结构发出各移动台的寻呼信号，并且使移动台从这些信号中识别出发给本台的信号呢？这就是多址接入方式要解决的问题。

5.1.2 多址方式

无线通信中，区分不同用户地址有多种方式，常见的多址方式有时分多址(TDMA)、频分多址(FDMA)、码分多址(CDMA)和空分多址(SDMA)。FDMA 中的多址连接信道是指各用户占用的频段；TDMA 中的多址连接信道是指各用户占用的时隙；CDMA 中的多址连接信道是指各用户占用的码型。图 5.1 分别给出了 N 个信道的 FDMA、TDMA 和 CDMA 的示意图。频分多址(FDMA)是以传输信号的载波频率不同来划分信道，一条宽带信道从频域分割为若干个子带，由滤波器实现多子带相对正交，即频带独享，时间共享。时分多址(TDMA)是以传输信号存在的时间不同来划分信道，每个不同信息流依序分占信道时隙，而每个时隙内的突发信息占用全部频带，即时隙独享，频率共享。码分多址(CDMA)是以传输信号的码型不同来划分信道，各用户都可以同时占用全部频带，实现正交信号传输，频率、时隙共享。空分多址(SDMA)是依据空间角度划分信道，多个在空间分割开来的信道或波束，可实现多个用户同时利用相同的频段通信，即频率、时隙、码型共享。由图 5.1 可见，频分多址中不同的用户的频道相互不重叠(正交)，时分多址中不同用户的时隙互不重叠，码分多址中不同用户的码型互不重叠。

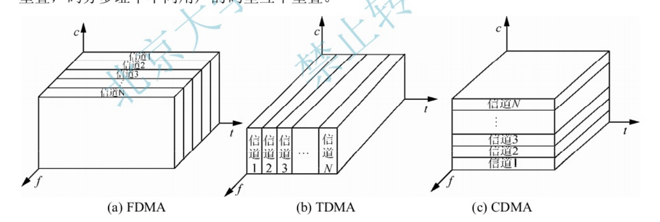


图 5.1 FDMA、TDMA 和 CDMA 的信道划分示意图

5.1.3 多址技术与多路复用的区别

多路复用是将来自不同信源的各路信息，按某种方式合并成一个多路信号，然后通过同一个信道传送给接收端。接收端再从该多路信号中按相应方式分离出各路信号，分送给不同的用户或终端。简而言之，多路复用是实现在同一信道中同时传输多路信号的一种技术，可以解决在同一信道内同时传送多个信号的问题。通过复用，可以把多路低速信号合并为一路高速信号进行传输，提高信道利用率。为了在接收端将不同路的信号区分开来，



必须使不同路的信号具有不同的特征。常见的多路复用方式有频分复用 FDM、时分复用 TDM、码分复用 CDM、波分复用 WDM 等。

多址技术是将多个用户的射频信号在射频信道上的复用，以实现多个用户同时共享一个通信信道，并独立地传输各自信息的通信技术。多址接入的方式有频分多址(FDMA)、时分多址(TDMA)和码分多址(CDMA)，还有利用不同地域区分用户的空分方式(SDMA)，利用正交极化区分的极化方式等。

多路复用和多址技术的相同点：多址接入和多路复用都是为了共享通信资源，理论基础都是信号的正交分割原理。

多路复用与多址技术的不同点如下。

(1) 多路复用一般在中频或基带实现，以达到两个台、站之间双边点对点的通信；多址技术通常在射频实现，以达到同一时间、同一方向的多用户间的多边通信。

(2) 多路复用中通信资源是预先分配给各用户；多址技术中通信资源通常是动态分配的，由用户在远端提出共享要求，在系统控制器控制下，按照用户对通信资源的需求，随时动态地改变通信资源的分配。

(3) 多路复用中多用户的要求是固定的，共享信道多半是在信道两端较集中区域进行多路接入，并集中提供复用和分路设备；多址技术是使多用户尽可能分散在任何地点，通过手机、移动台、车载台等设备，以随机接入方式复用信道，是实际接入与复用于一体的开放型多用户通信方式。

5.2 频分多址技术

5.2.1 FDMA 的基本原理

频分多址技术(FDMA)是将通信系统的总频段划分成若干等间隔的频道，并按要求分配给请求服务的不同用户使用。这些频道之间互不重叠，其宽度应能传输一路用户数据信息，而相邻频道之间应无明显的串扰，在呼叫的整个过程中，其他用户不能共享这一频段。频分多址是在不同频率上安排不同数据的多址方式，这有些类似于男女二重唱，在男女二重唱中男女声音的频率是不一样的，女人声调频率比男人的高，同一时间我们很容易根据频率的不同区分出男人和女人各自唱的内容。

频分多址以频率来分割信道，多个频道在频率轴上严格分开，但在时间上和空间上是重叠的，此时信道的含义即为频道。在移动通信系统中，每个频道宽度能传输一路话音信息。FDMA 系统工作示意图如图 5.2 所示，为了实现双工通信，收发双方使用不同的频率，称之为频分双工(Frequency Division Duplex，FDD)。在频分双工系统中，分配给用户两个频道，即接收和发送信道的频率是不同的，其中前向信道(基站向移动台方向的信道)频率较高，反向信道(移动台向基站方向的信道)频率较低。通信系统的基站必须同时发射和接收多个不同频率的信号，任意两个移动用户之间进行通信都必须经过基站的中转，因而必须同时占用 2 个信道(1 对频道)才能实现一对用户的双工通信。FDMA 系统中移动台在通信时所占用的频道并不是固定指配的，而是在通信建立阶段由系统控制中心临时分配的，通信结束后，移动台将释放它所占用的频道，这些频道又可以重新分配给别的用户使用。

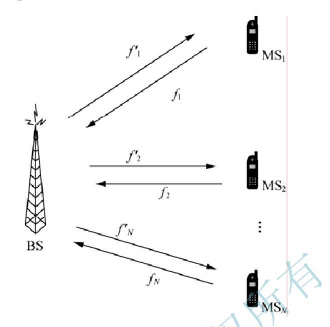


图 5.2 FDMA 系统的工作示意图

FDMA 系统频谱分割示意图如图 5.3 所示。由图可见，在频率轴上，前向信道占有较高的频带，反向信道占有较低的频带，两者之间留有保护频段，保护频段一般必须大于一定数值。此外，用户信道之间通常要设有载频间隔 Δf ，以避免系统频率漂移造成频道间的重叠。例如，在 GSM 900MHz 频段，收发频率间隔通常为 45MHz，载频间隔为 200kHz。

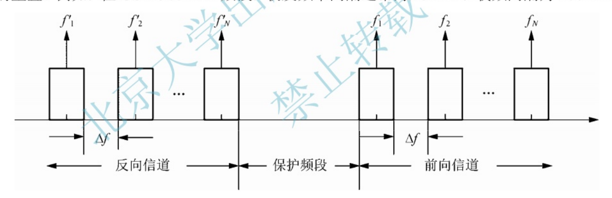


图 5.3 FDMA 系统频谱分隔示意图

5.2.2 FDMA 系统中的干扰问题

FDMA 系统是基于频率划分信道，每个用户在一对频道中通信，若有其他信号的成分落入一个用户接收机的频道带内时，将会造成对有用信号的干扰。FDMA 系统的主要干扰有互调干扰、邻道干扰和同频干扰。互调干扰是指系统内由于非线性器件产生的各种组合频率成分落入本频道接收机通带内造成对有用信号的干扰。邻道干扰是指相邻波道信号中存在的寄生辐射落入本频道接收机通带内造成对有用信号的干扰。同频干扰是在频率集重复使用的蜂窝系统中，频率相同信道的信号造成的干扰。

解决 FDMA 系统中这几种干扰的办法如下。



(1) 为了克服互调干扰，应尽可能提高系统的线性程度，减少发射机互调和接收机互调干扰，选用无互调的频率集。

(2) 为了克服邻道干扰，应严格规定收发信机的技术指标，限制规定发射机寄生辐射和接收机中频选择性，加大频道间的隔离度，这也涉及 FDMA 系统的频率规划问题。

(3) 为了减少同频道干扰，需要合理地选定蜂窝结构与频率规划，表现为系统设计中对同频道干扰因子的选择。

5.2.3　FDMA 的特点

FDMA 的特点有以下几方面。

(1) 单路单载频传输。FDMA 系统中每个信道的每个载波仅支持一路业务信息传输，信道的相对带宽较窄。

(2) 信号连续传输。系统分配给移动台和基站一对 FDMA 信道，在分配好话音信道后，基站和移动台直到通信结束都是连续传输的，然而在越区切换时，必须瞬时中断，以便将通信从一个频率切换到另一频率。对于语音，瞬时中断问题不大，对于数据传输则将造成数据的丢失。

(3) 频道受限和干扰受限。FDMA 系统载波带宽与单个信道一一对应，在接收设备中必须使用带通滤波器允许指定信道里的信号通过，滤除其他频率的信号，从而抑制了干扰。

(4) 基站复杂庞大，重复设置收发信设备。基站有多少信道，就需要多少部收发信机，同时需用天线共用器，功率损耗大，易产生信道间的互调干扰。

(5) 无须自适应均衡。FDMA 方式中，每个信道只传送一路数字信号，信号速率低，码元宽度与平均延迟扩展相比较是很大的，在数字信号传输中，由码间干扰引起的误码极小，因此在窄带 FDMA 系统中无须自适应均衡。

(6) 频率利用率低，容量小。FDMA 以频道来区分用户，频率资源确定了系统容量的大小。

在通信领域，FDMA 多址方式是最经典的多址技术，也是应用最早的一种多址技术。在卫星通信和各种无线通信网中，很多都采用频分多址，如美国的 AMPS、欧洲的 TACS、NMT-450、NMT-900、日本的 NTT 等移动通信系统采用的均是 FDMA 多址技术。第一个美国模拟蜂窝移动通信系统 AMPS 就是以 FDMA/FDD 为基础的，呼叫时，一个用户占用一个双工信道，这一信道实际上由两个具有 45MHz 间隔的单工信道构成。当一个呼叫完成或一个切换发生时，信道就空闲出来以便其他移动用户使用。

在 FDMA 系统中可以同时支持的信道数由下式计算

$$N=\frac{B_s-2B_p}{B_c} \tag{5-4}$$

其中 B_s 为系统带宽，B_p 为分配频谱时的保护带宽，B_c 信道带宽。为了避免工作在频带边缘的用户因为“溢出”而进入邻近的无线业务系统，需要在所分配的系统带宽的低端和高端设置保护带宽。

【例 5.1】如果 B_s 是 12.5MHz，B_p 是 10kHz，B_c 是 30kHz，求 FDMA 系统中的有效信道数。

解：将所给参数代入式(5-4)可求得在该 FDMA 系统中有效信道数为

$$N = \frac{12.5\times10^6 - 2(10\times10^3)}{30\times10^3} = 416$$

在我国，频分多址移动通信的频段是 890～905MHz(移动台发到基站)和 935～950MHz(基站发到移动台)，共 15MHz 的带宽，收发双工间隔 45MHz。如果按信道间隔为 25kHz 则可分为 600 个信道。

5.3 时分多址技术

5.3.1 TDMA 的基本原理

时分多址技术(TDMA)就是在不同的时间安排不同数据的多址方式，这类似于在大学教室的使用情况，如同一个教室，在不同时间安排不同的上课班级，而同一时间只能由特定的班级使用。TDMA 是以时间作为信号分割的参量来区分信道，它将载波在时间上分割成周期性的帧，每一帧再分割成若干时隙，每个时隙就作为一个通信信道分配给一个用户。帧与帧之间或时隙之间在时间上都是互不重叠的，系统根据一定的时隙分配原则，使各个移动台在每帧内只能按指定的时隙向基站发送信号。在满足定时和同步的条件下，基站可以分别在各时隙中接收到各移动台的信号而互不干扰。同时，基站发向各个移动台的信号都按顺序安排在预定的时隙中传输，各移动台只要在指定的时隙内接收，就能从时分多路复用(TDM)的信号接收到发给它的信号。图 5.4 是 TDMA 通信系统的工作示意图。

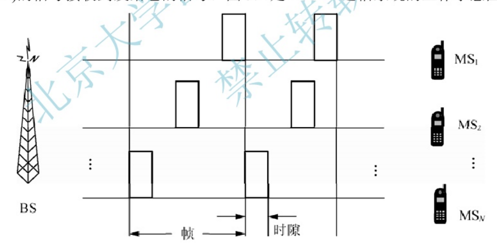


图 5.4 TDMA 系统的工作示意图

TDMA 结构较复杂，其最简单的情况就是单路载波被划分成许多不同的时隙，每个时隙传输一路突发信号。用户部分是 TDMA 的关键，每个用户在呼叫开始时被分配给一个时隙，用户与基站之间进行同步通信，并对时隙进行计数。当自己的时隙到来时，手机就启动接收和解调功能，对基站发来的突发信号进行接收。同样，当用户要发送信息时，首先是将信息进行缓存，等待自己的时隙到来，在时隙开始后，再将信息以加倍的速率发送出去，然后又开始积累下一次突发式传输。



TDMA 系统可以采用时分双工(TDD)或频分双工(FDD)两种工作方式(TDMA/TDD 和 TDMA/FDD)，这两种方式的 TDMA 系统分别如图 5.5 和图 5.6 所示。采用时分方式时，前向信道与反向信道的载波频率相同，因此将帧中的时隙分为两部分，一半用于前向信道，另一半用于反向信道，交替转换。采用频分双工方式时，前向信道与反向信道的载波频率不同，而且两个频率间要有一定的保护间隔。

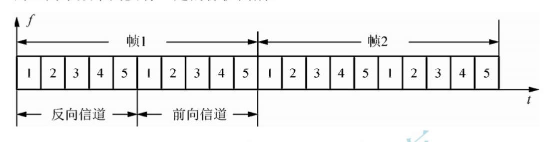


图 5.5　时分双工 TDMA 系统

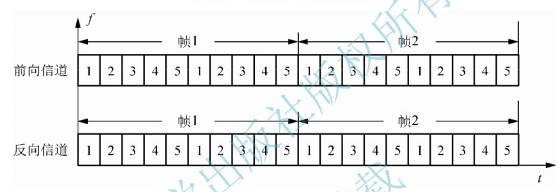


图 5.6　频分双工 TDMA 系统

5.3.2　TDMA 的帧结构

TDMA 帧是 TDMA 系统接收、处理和传输信息的基本单元，它由若干时隙组成。不同的用户周期性地占有一系列特定的时隙来传送自己的信息，但在每一个 TDMA 帧中，一个用户只占有一个特定时隙。不同通信系统的帧长度和帧结构是不一样的，典型的帧长在几毫秒到几十毫秒之间，例如 GSM 系统中每帧包含 8 个时隙，帧长为 4.6ms，DECT 系统中每帧有 24 个时隙，帧长为 10ms。

TDMA 的帧结构如图 5.7 所示，可以看出，一个 TDMA 帧是由头比特、信息数据和保护比特组成的。头比特包含了基站和用户用来确认彼此的地址信息和同步信息，信息段包含了各用户时隙，而保护比特则用来保证帧与帧之间的同步。在 TDMA 系统中，每帧中时隙结构的设计通常要考虑 3 个主要问题：一是控制和信令信息的传输；二是多径衰落信道的影响；三是系统的同步。

TDMA 系统正常工作的前提是同步和定时。通信双方只允许在规定的时隙内发送和接收信号，因此 TDMA 系统必须保证严格的帧同步、位同步、时隙同步和载波同步。TDMA 系统常采用的同步方式有位同步和帧同步两种。位同步是接收机正确解调的基础，用于传输位同步信息的方法一般有两种：一种是用专门的信道传输；另一种是插入业务信道中传输，如在每一个时隙的前面发送一段“0”、“1”交替的信号作为位同步信息。帧同步和位

同步常常采用相同的方法，即在每帧和每时隙前面分别设置一个同步码作为同步信息。同步码的选择是在帧长度确定后，根据信道条件和对同步的要求而确定的。对帧同步和位同步的要求是建立时间短、同步保持时间长、失步概率小等。

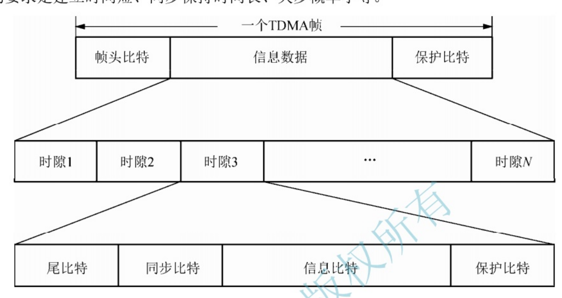


图 5.7　TDMA 的帧结构

系统定时是 TDMA 系统中的关键问题，系统常采用的定时方法有主从同步和独立时钟同步两种。主从同步是指系统所有设备的时钟均直接或间接地从属于某一个主时钟信息，主时钟通常有很高的精度，其信息以广播的方式传送给全网的很多设备，或者以分层的方式逐层传送给全网的设备，各设备从接收到的时钟信号中提取定时信息，并锁定到主时钟上。独立时钟同步是在网中各设备内均设置高精度的时钟，在通信开始或进行过程中，只要根据某一标准时钟进行一次时差校正，在很长时间内，时钟都不发生明显漂移，从而得到准确的定时。

5.3.3　TDMA 的系统容量和帧效率

1. 系统容量(用户数)

实际中，通常将 TDMA 方式与 FDMA 方式结合起来使用，这样可以大大提高系统的容量。TDMA 系统容量(用户数)的计算公式为

$$N=\frac{m\left(B_{\mathrm{s}}-2B_{\mathrm{p}}\right)}{B_{\mathrm{c}}} \tag{5-5}$$

其中 m 为每个频率信道所含的时隙数(TDMA 用户数)，B_{s} 为系统带宽，B_{p} 为保护带宽，B_{c} 信道带宽。

2. 帧效率

TDMA 系统的帧效率是指一帧中用户数据比特数占总比特数的百分比，计算公式为

$$\eta_f=\frac{n_{\mathrm{u}}}{n}\times100\%=\left(1-\frac{n_{\mathrm{s}}}{n}\right)\times100\% \tag{5-6}$$



其中 n_{u} 为用户数据的比特数，n 为每一帧的总比特数，n_{s} 为每一帧中系统开销的比特数，且 $n = n_{\mathrm{u}} + n_{\mathrm{s}}$。

【例 5.2】GSM 是一个前向信道占据 25MHz 频谱带宽的 TDMA/FDD 系统，该 25MHz 频谱带宽被分成若干带宽为 200kHz 的无线信道。如果一个无线信道支持 8 个语音信道，并且假设没有保护带宽，那么 GSM 中能够同时容纳的用户数为多少？

解：频谱带宽 B_{s}=25MHz，信道带宽 B_{c}=200kHz，由于没有保护带宽，即 B_{p}=0，又由于每一个无线信道能支持 8 个语音信道，即可同时支持 8 个用户，m=8，则 GSM 系统能够同时容纳的用户数为

$$N = \frac{m(B_{\mathrm{s}} - 2B_{\mathrm{p}})}{B_{\mathrm{c}}} = \frac{8 \times 25\mathrm{MHz}}{200\mathrm{kHz}} = 1000$$

【例 5.3】考虑每帧支持 8 个用户且数据速率为 270.833Kbps 的 GSM/TDMA 系统，试求：

(1) 每一用户的原始数据速率是多少？

(2) 在保护时间、跳变时间和同步比特共占用 10.1Kbps 的情况下，每一个用户的传输效率是多少？

解：(1) 每用户的原始数据速率：270.833/8 = 33.854(Kbps)

(2) 传输效率：$\eta_f = \left(1 - \frac{n_{\mathrm{s}}}{n}\right) \times 100\% = \left(1 - \frac{10.1}{33.853}\right) \times 100\% = 70\%$

5.3.4　TDMA 的特点

TDMA 的特点有以下几方面。

(1) TDMA 系统分配给每个用户一个互不重叠的时隙，各用户仅在所分配的时隙工作，N 个用户可以共享一个频带资源，所以它的频带利用率高，系统容量大，互调干扰小。

(2) TDMA 系统 N 个时分信道共用一个载波频道，占据相同带宽，只需一部收发信机，所以基站复杂性较小。

(3) TDMA 用不同的时隙来发送和接收信号，因此不需要双工器。

(4) TDMA 用户工作在非连续的突发状态，电源效率高，TDMA 系统还可以根据用户需求灵活地进行时隙分配，必要时可以将多个时隙分配给一个用户使用。

(5) TDMA 系统的发射响应速率较高，为了消除码间干扰的影响需要采用自适应均衡，用以补偿传输失真。

(6) 同步技术是 TDMA 系统正常工作的重要保证，TDMA 用于同步控制等的系统开销较大。

(7) 越区切换简单。由于在 TDMA 中移动台是不连续的突发式传输，所以切换处理对一个用户单元来说是很简单的，它可以利用空闲时隙监测其他基站，这样越区切换可在无信息传输时进行，因而没有必要中断信息的传输，传输数据不会因越区切换而丢失。

在移动通信中最典型的时分多址方式有北美的 D-AMPS、欧洲与我国的 GSM-900、DCS-1800、日本的 PDC。欧洲的 GSM 系统采用 TDMA/FDMA 技术，系统有多个载频，每个载频上再分成 8 个时隙。美国的 D-AMPS 数字蜂窝标准采用 FDMA/TDMA 相结合，采用 30kHz 的 FDMA 信道，并将每个信道再分割成 6 个时隙，用于 TDMA 传输。日本的 PDC 系统采用 TDMA/FDMA 技术。

5.4 码分多址技术

5.4.1 CDMA 的基本原理

码分多址技术(CDMA)以码型结构作为信号分割的参量，它为每个用户分配了各自特定的地址码，利用公共信道来传输信息，接收端通过地址码的不同来识别不同的用户。码分多址是类似于人们在国际会议中用不同的语言如英语、汉语等交流信息的多址方式，在同一时间，同样的频率，只要说话的人和听话的人使用相同的语言就可以彼此交谈。在无线通信中，对信息内容进行编码的不是语言，而是不同的地址码。CDMA 系统的地址码相互正交，主要用于区别不同地址，地址码在频率、时间和空间上都可能重叠，对某用户发送的信号，系统的接收端需用完全一致的本地地址码对接收的信号进行相关检测。只有与其相匹配的接收机通过相关检测才可能正确接收，其他使用不同码型的信号因为和接收机本地产生的码型不同而不能被解调。

CDMA 系统的工作原理示意图如图 5.8 所示。在 CDMA 移动通信系统中，用户之间的信息传输也是由基站进行转发和控制的，不同的移动用户传输信息所用的信号不是靠频率不同或时隙不同来区分的，而是用各自不同的编码序列来区分的。为了实现双工通信，CDMA 系统可以采用频分双工和时分双工工作方式。如在频分双工码分多址工作方式中(图 5.9)，每一个用户分配一个地址码，这些码型互不重叠即相互正交，并在同一载波上进行传输，其用户的前向信道和反向信道采用频率划分从而实现双工通信。

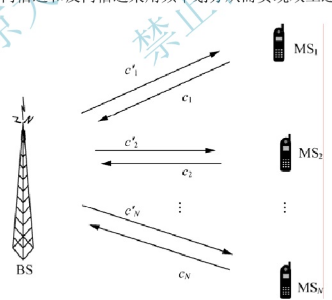


图 5.8 CDMA 系统的工作原理示意图

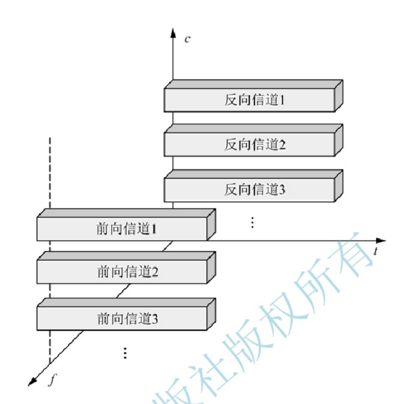


图 5.9　频分双工 CDMA 系统

根据扩频实现方法的不同，码分多址可以分为直接序列码分多址(Direct Sequence CDMA，DS-CDMA)、跳频码分多址(Frequency Hopping CDMA，FH-CDMA)、跳时扩频码分多址(Time Hopping CDMA，TH-CDMA)，也可以将上述基本方法进行组合从而形成各种混合码分多址技术(HCDMA)。

直接序列码分多址(DS-CDMA)是目前用得较多的一种扩频多址方式，它在第二代移动通信中已获得了成功的应用，并且也是第三代移动通信的核心技术。DS-CDMA 是通过将携带信息的窄带信号与高速的地址码信号相乘而获得的宽带扩频信号。通过 DS 扩频，将信号功率谱在一个很宽的频谱上进行了“平均”，或者说是在背景噪声不变的情况下，信噪比 S/N 变得很低，好像是将信号在噪声中“隐藏”了起来。因此 DS-CDMA 系统具有抗窄带干扰、抗多径衰落和保密性好等优点。

在跳频码分多址系统中，每个用户根据各自的伪随机(PN)序列，动态地改变已调信号的中心频率。在信息传输过程中，发送设备不断更换频率。接收设备必须使用同样的跳频规律才能正确解扩。FH-CDMA 类似于 FDMA，但 FH-CDMA 所用的频道是动态变化的。FH-CDMA 通信过程中要求频率序列正交，在一个 PN 序列周期内，所有使用频率在任意时刻必不相同。

与跳频码分多址不同的是，跳时码分多址用一组正交跳时码控制各个用户的通信信号在一帧时间内的不同位置进行伪随机跳变，跳时码分多址 TH-CDMA 可以看成是一种由伪



随机码控制的多进制脉冲位置调制(MPPM)。为了进一步提高抗干扰性能，TH-CDMA 通常都是与其他扩频技术(如跳频)混合使用的。

混合码分多址(HCDMA)是指码分多址之间或者码分多址与其他多址方式之间混合使用的多址方式，混合码分多址可以克服使用单一多址方式的缺点，进而获得优势互补的效果。组合的具体方式有多种，常用的混合码分多址方式有跳频与跳时结合的 FH/TH-CDMA、跳频与直接序列相结合的 FH/DS-CDMA、跳时与直接序列相结合的 TH/DS-CDMA；码分多址与其他多址方式的组合形式有：FDMA 与 DS-CDMA 相结合的 FD/DS-CDMA、TDMA 与 DS-CDMA 相结合的 TD/DS-CDMA 以及 TDMA 与 FH-CDMA 相结合的 TD/FH-CDMA 等。

5.4.2 伪随机码

CDMA 码分多址技术给每个用户分配一个唯一的地址码，通过该地址码的不同来识别用户，因此地址码的选择直接影响 CDMA 系统的容量、抗干扰能力、接入和切换速度等。CDMA 系统中的地址码比特率应能满足扩展带宽的需要；应具有尖锐的自相关特性，且互相关要小；应具有近似噪声的频谱性质，即近似连续谱，且均匀分布。为提高抗干扰能力，地址码通常使用各种伪随机码(Pseudo-Noise，PN)，也称伪码，又称伪随机序列。伪随机码的码型影响码序列的相关性，其码元长度决定扩展频谱的宽度。常用的伪随机码有 m 序列、Gold 序列和 Walsh(沃尔什)码。下面我们简单介绍 Walsh 码、m 序列和 Gold 序列。

1. Walsh 码

Walsh 函数又称 Walsh 码，相应的离散 Walsh 函数简称为 Walsh 码序列。Walsh 函数集是完备的非正弦正交函数集，有着处处为零的互相关(正交性)和较好的自相关特性。Walsh 函数波形如图 5.10 所示，下面就以该波形为例说明 Walsh 函数的特性。由图 5.10 可知，Walsh 码序列的波形是定义在半开区间 $[0,1)$ 的矩形波族。设 $n(n=0,1,2,\cdots)$ 是矩形波的编号，将编号为 n 的 Walsh 码表示为 $\mathrm{Wal}(n,t)$，矩形波幅度的取值仅有+1 或−1 两种，因此也比较适合用来表示和处理数字信号。若规定起始波形的取值为+1，则之后的 Walsh 码序列的波形在+1 与−1 之间变化。通常，信息“0”用 Walsh 函数的正向波形表示，信息“1”用 Walsh 函数的反向波形表示，数字信息通过这样的调制后再进行传输。

Walsh 函数最重要的性质就是正交性，即

$$\int_0^1 \mathrm{Wal}(i,t)\cdot\mathrm{Wal}(j,t)\mathrm{d}t=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}\quad (i,j=0,1,2,\cdots) \tag{5-7}$$

利用 Walsh 函数的正交性，可获得 CDMA 的地址码。Walsh 函数在同步时是完全正交的，只有用户的地址码波形与接收到的信号波形码结构完全相同时，才能从中获取有用信息。

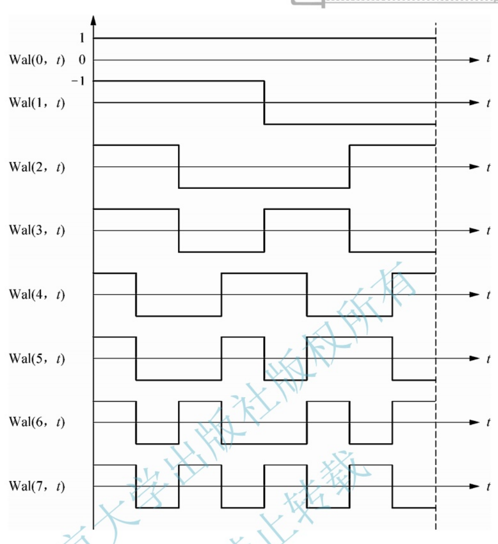


图 5.10　Walsh 函数的波形

对图 5.10 中的 Walsh 函数波形在 8 个等间隔上取样，就得到离散 Walsh 函数，图中所示的 Walsh 函数可用 8×8 的 Walsh 函数矩阵表示

$$\begin{bmatrix} 00 & 00 & 00 & 00 \\ 00 & 00 & 11 & 11 \\ 00 & 11 & 11 & 00 \\ 00 & 11 & 00 & 11 \\ 01 & 10 & 01 & 10 \\ 01 & 10 & 10 & 01 \\ 01 & 01 & 10 & 10 \\ 01 & 01 & 01 & 01 \end{bmatrix} \tag{5-8}$$

Walsh 函数通常用哈达玛(Hadamard)矩阵 $\boldsymbol{H}$ 表示，哈达玛矩阵 $\boldsymbol{H}$ 是由元素“1”和“0”构成的正交方阵，利用递推关系很容易构成 Walsh 函数序列族。在哈达玛矩阵中，任意两行(列)都是正交的，即互相关函数为零，因此，将 M 阶哈达玛矩阵中的每一行定义为一个 Walsh 码序列时，就得到 M 个 Walsh 序列。Walsh 函数矩阵(哈达玛矩阵)的递推关系如下

$$H_0=(0) \qquad H_2=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$$

$$H_4=H_{2\times2}=\begin{bmatrix}H_2 & H_2\\ H_2 & \overline{H}_2\end{bmatrix}=\begin{bmatrix}00 & 00\\ 01 & 01\\ 00 & 11\\ 01 & 10\end{bmatrix}$$

$$H_8=H_{4\times4}=\begin{bmatrix}H_4 & H_4\\ H_4 & \overline{H}_4\end{bmatrix}=\begin{bmatrix}00 & 00 & 00 & 00\\ 00 & 00 & 11 & 11\\ 00 & 11 & 11 & 00\\ 00 & 11 & 00 & 11\\ 01 & 10 & 01 & 10\\ 01 & 10 & 10 & 01\\ 01 & 01 & 10 & 10\\ 01 & 01 & 01 & 01\end{bmatrix} \tag{5-9}$$

……

$$H_{2M}=\begin{bmatrix}H_M & H_M\\ H_M & \overline{H}_M\end{bmatrix}$$

其中 M 是 2 的整数次幂，$\overline{H}_M$ 是 H_M 的补。

Walsh 码序列的自相关特性并不理想，但互相关特性很好，两两之间的互相关函数为 0，因而在码分多址通信中，Walsh 码序列可以作为地址码使用。在 IS-95 中，正向信道使用 64 阶 Walsh 码序列主要是用于复用目的，用来区分各信道，而反向信道使用 64 阶 Walsh 码序列则是用作正交码。在第三代移动通信 WCDMA、CDMA2000 和 TD-SCDMA 系统中也都是以 Walsh 码来划分信道的。

2. m 序列伪随机码

m 序列是最长线性移位寄存器序列的简称，它是最基本、最重要的一种伪随机序列。m 序列规律性强、易于产生、有许多优良特性，在码分多址系统中应用最早。目前的 CDMA 蜂窝移动通信系统中就是采用 m 序列作为地址码，并利用它的不同相位来区分不同的用户。m 序列是由带线性反馈的移位寄存器产生的周期最长的一种序列，它的周期是 $P=2^n-1$，n 是移位寄存器的级数。m 序列发生器由移位寄存器、反馈抽头及模 2 加法器组成，图 5.11 是一个由四级移位寄存器构成的 m 序列发生器。

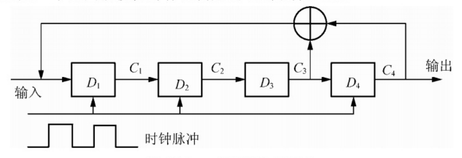


图 5.11 m 序列发生器模型



m 序列有许多优良的特性，下面主要讨论 m 序列的随机性、自相关性和互相关性。

1) m 序列的随机性

m 序列的随机性主要有以下几方面。

(1) 均衡性。m 序列一个周期内“1”和“0”的码元数大致相等，其中“1”比“0”只多一个，这个特性保证了在扩频系统中，用 m 序列作平衡调制实现扩展频谱时有较高的载波抑制度。

(2) 游程分布。m 序列中连续“0”或“1”的元素称为游程，一个周期内($P=2^n-1$)长度为 k ($0\leqslant k\leqslant n-2$) 的游程中，连“1”的游程和连“0”的游程各占一半，而且只有一个包含 $n-1$ 个“0”的游程，也只有一个包含 n 个“1”的游程。

(3) 移位相加特性。m 序列和其移位后的序列逐位模 2 和，所得的序列还是 m 序列，只是相位不同。

(4) m 序列发生器中的移位寄存器的各种状态，除全 0 外，其他状态在一个周期内只出现一次。

2) m 序列的自相关性

m 序列的自相关性用自相关函数来描述，一个周期为 $P=2^n-1$ 的 m 序列 $\{a_n\}$ (a_n 取值为 1 或 0)与其左移 τ 位的序列 $\{a_{n+\tau}\}$ 的自相关函数表示为

$$R_a(\tau)=\frac{A-D}{A+D} \tag{5-10}$$

其中 A 为 $\{a_n\}$ 和 $\{a_{n+\tau}\}$ 逐位模 2 和后的序列中“0”的位数或序列 $\{a_n\}$ 和 $\{a_{n+\tau}\}$ 相同的位数；D 为“1”的位数或序列 $\{a_n\}$ 和 $\{a_{n+\tau}\}$ 不相同的位数，$P=A+D$ 。

如果伪随机码的码元用 1 和−1 表示，与 0 和 1 表示法的对应关系是“0”变成“1”，“1”变成“−1”，即 m 序列中 a_n 的取值是−1 和 1，此时自相关函数表示为

$$R_a(\tau)=\frac{1}{P}\sum_{n=1}^{P}a_n a_{n+\tau}=\begin{cases}1, & \tau=0\\ -1/p, & \tau\neq 0\end{cases} \tag{5-11}$$

图 5.12 为 m 序列自相关函数图。由图可知，当 $\tau=0$ 时，m 序列的自相关函数 $R(\tau)$ 出现峰值 1；当 τ 偏离 0 时，相关函数曲线很快下降；当 $1\leqslant\tau\leqslant P-1$，相关函数值为 $-1/P$；当时 $\tau=P$，又出现峰值，如此周而复始。当周期 P 很大时，m 序列的自相关函数与白噪声类似，相关检测就是利用这一特性检测自相关函数值为 1 的码序列。

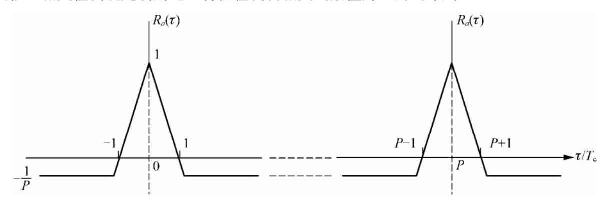


图 5.12　m 序列自相关函数图

【例 5.4】 m 序列$\{a_n\}$=100011110101100，求$R_a(0)$，$R_a(1)$，…，$R_a(14)$，画出自相关函数曲线。

解： $\tau=0$时，$\{a_{n+0}\}$=100011110101100，$A=15$，$D=0$，$R_a(0)=1$

$\tau=1$时，$\{a_{n+1}\}$=000111101011001，$A=7$，$D=8$，$R_a(1)=-1/15$

同理，$R_a(2)=\cdots=R_a(14)=-1/15$。自相关函数曲线如图 5.13 所示。



图 5.13　m 序列的自相关函数曲线

3) m 序列的互相关性

m 序列的互相关性是指周期相同的两个不同 m 序列$\{a_n\}$和$\{b_n\}$一致性的程度，互相关值越接近 0，说明这两个 m 序列差别越大，互相关性越弱，反之，说明这两个 m 序列差别较小，互相关性较强。当 m 序列用作 CDMA 系统的地址码时，为了避免用户之间的相互干扰，必须选择互相关值较小的 m 序列组。

周期为$P=2^n-1$的 m 序列$\{a_n\}$与其左移τ位的序列$\{b_{n+\tau}\}$(a_n、b_n取值为 1 或 0)的互相关函数表示为

$$R_c(\tau)=\frac{A-D}{A+D} \tag{5-12}$$

其中A为$\{a_n\}$和$\{b_{n+\tau}\}$逐位模 2 和后的序列中“0”的位数或序列$\{a_n\}$和$\{b_{n+\tau}\}$相同的位数，D为“1”的位数或序列$\{a_n\}$和$\{b_{n+\tau}\}$不相同的位数，$P=A+D$。

同样，如果伪随机码的码元用 1 和-1 表示，此时这两个不同 m 序列$\{a_n\}$和$\{b_{n+\tau}\}$的互相关函数表示为

$$R_c(\tau)=\frac{1}{P}\sum_{n=1}^{P}a_nb_{n+\tau} \tag{5-13}$$

周期相同的 m 序列组中，两两 m 序列对的互相关性差异较大。在实际应用中，通常只关心互相关特性较好的 m 序列对。对于周期为$P=2^n-1$的 m 序列组，其最好的 m 序列对的互相关函数只有 3 个值，分别为

$$R_c(\tau)=\begin{cases}\dfrac{t(n)-2}{P}\\-\dfrac{1}{P}\\-\dfrac{t(n)}{P}\end{cases} \tag{5-14}$$



式中，$t(n)=1+2^{[(n+2)/2]}$，其中[ ]表示取实数的整数部分。这 3 个值被称为理想三值，能够满足这一特性的 m 序列对就是 m 序列优选对。

在 CDMA 数字蜂窝移动通信系统中，可为每个基站分配一个伪随机序列，以不同的伪随机序列来区分基站地址，也可以只用一个伪随机序列，而用伪随机序列的相位来区分基站地址，即每个基站分配一个伪随机序列的初始相位。IS-95 CDMA 数字蜂窝移动通信系统就是采用给每个基站分配一个伪随机序列的初始相位的方法，利用了 m 序列良好的自相关特性，采用的伪随机序列周期为 $2^{42}-1$ 的长码来区分用户。

3. Gold 序列

R.Gold 于 1967 年提出了一种基于 m 序列优选对的码序列，称为 Gold 序列。Gold 序列是 m 序列的组合码，它由两个码长相等、码时钟速率相同的 m 序列优选对逐位模 2 和得到，当改变其中一个 m 序列的相位时，可得到一个新的 Gold 序列。Gold 序列具有优良的自相关和互相关特性、构造简单、产生的序列数多，常常用于码分多址中区分基站和用户。Gold 序列的生成电路如图 5.14 所示，图中 m 序列发生器 1 和 2 产生的 m 序列是一个 m 序列优选对，m 序列发生器 1 的初始状态固定不变，调整 m 序列发生器 2 的初始状态，在同一时钟脉冲控制下，生成的两个 m 序列经过模 2 和后可得到 Gold 序列。通过设置 m 序列发生器 2 的不同初始状态，可以得到不同的 Gold 序列。

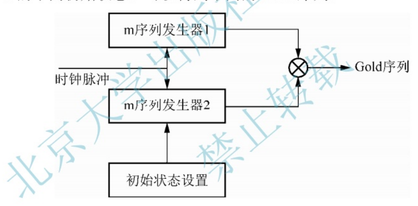


图 5.14　Gold 序列的生成

实际工程中，主要考虑 Gold 序列的 3 个特性：相关特性、Gold 序列的数量和平衡的 Gold 序列。

相关特性：具有与 m 序列优选对相类同的自相关和互相关特性。当 $\tau=0$ 时，Gold 序列的自相关函数与 m 序列相同；当 $1\leqslant\tau\leqslant P-1$ 时，自相关函数取式(5-14)中 3 个理想值中的第三个值。

Gold 序列的数量：周期为 $P=2^n-1$ 的 m 序列优选对生成的 Gold 序列总共有 2^n+1 个，随着 n 的增加，Gold 序列数以 2 的 n 次幂增长，因此 Gold 序列数比 m 序列数多得多。

平衡的 Gold 序列：平衡的 Gold 序列是指在一个周期内“1”码元数比“0”码元数仅多一个。对于周期为 $P=2^n-1$ 的 m 序列优选对生成的 Gold 序列，当 n 是奇数时，有 $2^{n-1}+1$ 个 Gold 序列是平衡的，约占 50%；当 n 是偶数(不是 4 的倍数)时，有 $2^{n-1}+2^{n-2}+1$ 个 Gold 序列是平衡的，约占 75%，其余的都是不平衡的 Gold 序列。只有平衡的 Gold 序列才可以用到码分多址通信系统中去。

Gold 序列良好的特性，使其在多址技术中，尤其是码序列长度较短情况下得到广泛应用。在第三代移动通信 WCDMA 和 CDMA2000 体制中均采用了 Gold 码序列来区分小区(或扇区)以及不同小区(或扇区)之间的用户。在 WCDMA 系统中，下行链路采用 Gold 码区分小区和用户，上行链路采用 Gold 码区分用户。

5.4.3 CDMA 系统存在的问题

1. 多址干扰

CDMA 系统的接收端必须用完全一致的本地码对接收信号进行相关检测，其他使用不同码型的信号因为与接收机本地产生的码型不同而不能被解调，但它们的存在类似于在信道中引入了噪声或干扰，这通常称为多址干扰(MAI)。多址干扰是由于不同用户地址码之间的非完全正交性而造成的，通信过程中不同用户的发射信号会相互干扰，在异步传输信道以及多径传播环境中多址干扰将更严重。由于这种干扰是系统本身产生的，因此也称自干扰。解决多址干扰的方法是找到同步状态下和非同步状态下序列的互相关系数均为零的数字序列。

2、远近效应

假定 CDMA 系统中所有的用户发送功率都相同，则离基站近的用户的接收功率就会高于离基站远的用户的接收功率，这样由于传输距离不同(即传输衰减不同)就会造成不同位置的用户的信号在基站的接收状况差异较大，近距离用户的强信号抑制了远距离用户的弱信号，降低有效用户数，即使各用户到基站距离相等，深衰落的存在也会使到达基站的信号各不相同，强信号对弱信号有着明显的抑制作用，会使弱信号的接收性能很差甚至无法通信，这就是 CDMA 系统的远近效应。为了削弱远近效应的影响，大多数 CDMA 系统采用自动功率控制技术。

5.4.4 CDMA 的容量

CDMA 的载波干扰比(C/I)定义式为

$$\frac{C}{I}=\frac{R_{\mathrm{b}}E_{\mathrm{b}}}{N_0W}=\frac{E_{\mathrm{b}}}{N_0}\Big/\frac{W}{R_{\mathrm{b}}} \tag{5-15}$$

其中 E_{b} 为信息的比特能量，R_{b} 为信息的比特速率，W 为总频段宽度，N_0 为干扰(包括噪声)的功率谱密度，E_{b}/N_0 为归一化信噪比，W/R_{b} 为系统的处理增益。

在单小区时，考虑上行链路的容量，假设理想的功率控制(即在上行链路对所有移动台的发射功率进行控制，使到达基站接收机的信号功率均相同)，则基站接收机的载波干扰比为

$$\frac{C}{I}=\frac{P_{\mathrm{r}}}{(m-1)P_{\mathrm{r}}}=\frac{1}{m-1} \tag{5-16}$$

其中 P_{r} 为基站接收到的每个移动台的信号功率，m 为同时工作的移动台数，则一个小区内最大用户数为

$$m=1+\frac{W/R_{\mathrm{b}}}{E_{\mathrm{b}}/N_0} \tag{5-17}$$



上式是未考虑背景热噪声的结果，如果考虑背景热噪声 η，则接入系统的用户数为

$$m = 1 + \frac{W/R_b}{E_b/N_0} - \frac{\eta}{C} \tag{5-18}$$

该式表明，在误比特率一定的条件下，降低热噪声功率，减小归一化信噪比，增大系统的处理增益都有利于提高系统容量。由于 CDMA 系统是一个干扰受限的系统，所以减小干扰会使 CDMA 系统的容量线性增加。

【例 5.5】 IS-95CDMA 系统中，设 $E_b/N_0 = 7\text{dB}$，信息比特率 $R_b = 9.6\text{Kbps}$，传输频道带宽为 1.25MHz，求此系统中每个小区可容纳的用户数。若传输频道带宽为 25MHz，则此系统中每个小区可容纳的用户数为多少？

解： 传输频道带宽为 W=1.25MHz 时，可容纳的用户数为

$$m = 1 + \frac{1.25\times10^6/9.6\times10^3}{10^{0.7}} = 27$$

传输频道带宽为 W=25MHz 时，可容纳的用户数

$$m = 1 + \frac{25\times10^6/9.6\times10^3}{10^{0.7}} = 540$$

5.4.5　CDMA 的特点

CDMA 的特点有以下几方面。

(1) 频率共享。CDMA 系统可以实现多用户在同一时间内使用同一频率进行各自的通信而不会相互干扰。

(2) 系统容量大。CDMA 是干扰受限系统，任何干扰的减少都会直接转化为系统容量的提高，所以可以利用一些抗干扰技术来提高系统容量。通过计算，在相同条件下，CDMA 的系统容量大约是 TDMA 系统的 5 倍，是 FDMA 系统的 12 倍。

(3) 具有软容量特性。TDMA 可同时接入的用户数是一定的，超出这个数量的用户只能被阻塞掉而无法接入，对于 CDMA 系统，当系统满载时，另外增加少量用户，只会使通信质量略有下降，不会出现硬阻塞的现象，这就是 CDMA 系统的软容量特性。这个软容量特性，使得 CDMA 系统对用户数目没有绝对限制。

(4) 抗多径衰落。由于 CDMA 是扩频系统，其信号被扩展在一个较宽的频谱上，从而可以减少多径衰落。

(5) 低信号功率谱密度。CDMA 系统中，信号功率被扩展到较宽频谱上，其功率谱密度大大降低，这样使信号具有较强的抗窄带干扰能力，并且对窄带系统的干扰很小，可以与其他系统共用频段，使有限的频谱资源得到更充分的利用。

(6) 信道数据速率很高，无需自适应均衡。CDMA 系统中由于伪随机序列有很好的自相关性，使大于一个码片宽度的时延扩展部分受到接收机的自然抑制，此外，采用分集接收最大比合并技术，可获得最佳的抗多径衰落效果。

(7) 有效的宏分集和平滑的软切换，保证通信的连续性。

阅读材料

手机辐射问题

手机已经成为人们生活中不可分割的一部分，而人们对于无线通信设备所发出的射频电磁场对健康可能产生的不利影响也越发警惕。当您用移动电话进行呼叫时，移动电话会发射电磁波(又称射频能量)，这些电磁波可被距离最近的基站接收。一旦基站接收到移动电话传来的电磁波，就会将其传输到交换台，交换台根据当前呼叫的类型将呼叫转接到另一个基站或固定电话线网络，从而实现通话。

当人们使用手机时，手机会向基站传送电磁波，而这种电磁波或多或少都会被人体吸收，从而可能改变人体的组织，对人体的健康带来影响，这些电磁波就被称为手机辐射。使用手机就避免不了手机辐射，为了让手机辐射维持在一个相对更低更安全的一个范围内，国际上通常使用“SAR”(Specific Absorption Rate)即“比吸收率”来对手机辐射进行量化和测量，比吸收率是指单位时间内单位质量的物质吸收的电磁辐射能量。SAR 值越低，辐射被吸收的量越少。目前国际通用的标准有两个，一个是美国标准 1.6W/kg，另一个是欧洲与中国标准 2W/kg，其中 2W/kg 的具体含义是指以任意 6 分钟记时平均，每公斤人体组织吸收的电磁辐射能量不得超过 2W。虽然手机辐射并不会对我们产生直接的严重病症，但对人体的危害还是存在的，因此我们要尽可能地避免过多的手机辐射。

减小手机辐射建议如下。

1. 在手机信号较好的环境下使用手机，避免在墙角或较为封闭的电梯、地铁等地方使用。较为封闭的环境中信号一般较弱，手机在使用时需要产生更大功率的电磁波来保证通话。

2. 接打电话时尽量使用有线耳机或者蓝牙，让手机远离头部，或者还可以开启免提等功能来避免头部与手机的近距离接触。

3. 手机接通及拨出电话前后的几秒钟是手机电磁辐射最强的时候，因此在这两个时间段，最好不要让手机贴身，或贴耳接听。

4. 连续长时间通话会加大辐射对脑部的影响，建议长时间通话使用座机电话或耳机。如果不得不长时间用手机通话，应左右耳轮流接听。

5. 避免通话时频繁移动位置，频繁移动位置会影响手机信号的强弱，从而导致手机功率增大而产生更大辐射。

6. 睡眠或不使用手机时，因将手机放置在离人体较远的地方，切勿放在枕边及裤兜，手机在待机状态下同样有辐射，只是比较小而已。

5.5 空分多址技术

5.5.1 SDMA 的基本原理

空分多址(SDMA)是利用空间位置的不同来划分信号的多址方式。SDMA 技术通过空间的分割来区别不同用户，它利用天线的方向性波束将小区划分为不同的子空间来实现空间的正交隔离。移动通信中的扇区天线可以看成是 SDMA 的一种基本实现方式，在蜂窝移动通信中由于充分运用了这种多址方式，依靠严格的功率控制，精确地安排频谱，有效地实现空分多址，极大地发挥有限频谱的效能，使有限的频谱能构成大容量的通信系统。



因此，空分多址 SDMA 也成为蜂窝移动通信中的一项关键技术。我国提出的第三代移动通信标准 TD-SCDMA 中用到的就是 SDMA 技术。

SDMA 利用用户不同的空间特征区分用户，目前利用最多也是最明显的特征就是用户的位置，根据电磁波传播的特征可以使不同区域的用户在同一时间使用相同频率实现互不干扰的通信。例如，可以利用定向天线或窄波束天线，使电磁波按一定指向辐射，可以控制波束范围，不同波束范围可以使用相同频率，也可以控制发射的功率，使电磁波只能作用在有限的距离内。定向窄波束辐射时的空分多址示意图如图 5.15 所示。SDMA 在相同时隙、相同频率或相同地址码的情况下，可以根据信号不同的空间传播路径进行区分。SDMA 使用不同的天线波束为不同区域的用户提供接入。实际上，由于空间的分割不可能太细，即一般情况下某一空间范围内不可能只有一个用户，所以 SDMA 通常都不是单独使用，而是与其他多址方式如 FDMA、TDMA 和 CDMA 等结合使用。也就是说，对于处于同一波束内的不同用户，再用其他多址方式加以区分。

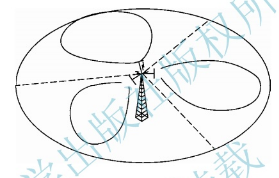


图 5.15　SDMA 系统示意图

传统的空分多址只是利用用户的位置特征，如何更充分地利用用户的空间特征，进一步提高空分多址的效能，是一个新的研究热点。近年来，随着技术的飞速发展，人们发现空间特征不仅是位置，一些当时认为无法利用的空间特征现在正逐步利用，形成以智能天线为基础的新一代空分多址方式，以自适应阵列处理为基础的智能天线技术应用于蜂窝系统为空分多址开辟了新的渠道。在移动通信中，采用自适应阵列天线是实现空间分割的基本技术，它可在不同用户方向上形成不同的波束。波束形成天线采用智能天线，基站的智能天线形成多个波束覆盖整个小区，智能天线可定位于每个移动终端。不同的波束可采用相同的频率和相同的多址方式，也可采用不同的频率和不同的多址方式。使用 SDMA 技术的智能自适应天线将是 4G 系统中一项新的关键技术，它可以实现最佳的 SDMA，具有无穷小波束宽度、无穷大快速搜索能力和无限快的跟踪速度，它能提供在本小区内不受其他用户干扰的唯一信道，还能克服多径干扰和同信道干扰问题。

5.5.2　SDMA 的特点

SDMA 的特点有以下几方面。

(1) SDMA 可以提高天线增益，使功率控制更加合理有效，显著地提高了系统容量。

(2) SDMA 可以削弱来自外界的干扰，同时降低对其他电子系统的干扰。

(3) SDMA 的关键技术是智能天线技术，无线信道的复杂性使得智能天线对多用户的动态捕捉、识别与跟踪以及信道的辨识等极为复杂。

(4) SDMA 在反向链路上实现起来较困难。其主要原因：第一，基站完全控制了在前向链路上所有发射信号的功率，但是由于每一用户和基站间无线传播路径不同，因此从每一用户单元出来的发射功率必须动态控制，以防止任何用户功率太高而干扰其他用户；第二，发射功率受到用户单元电池能量的限制，因此也限制了反向链路上对功率的控制程度。

5.6　正交频分复用多址技术

正交频分复用(OFDM)是一种调制方式，它可以很容易地与传统的多址技术结合实现多用户接入服务，如 OFDM-TDMA、OFDMA 和多载波 CDMA 等，本节将介绍这些多址技术。

5.6.1　OFDM-TDMA

正交频分复用时分多址方式(OFDM-TDMA)是一种以时间来区分用户的多址接入方式。在 OFDM-TDMA 系统中，信息的传输是按时域上的帧来进行的，每个 TDMA 帧包含多个时隙，每个时隙的宽度等于一个 OFDM 符号的时间长度，有信息要传送的用户按各自的需求可以占用一个或多个 OFDM 符号。每个用户在信息传输期间，将占用所有的系统带宽，即该用户的信息可以在 OFDM 的所有子载波上进行分配。在 OFDM-TDMA 系统中，可以使用自适应调制技术，使各个载波的调制方式不同，调制方式的选择主要是依据子载波的信噪比。在 OFDM-TDMA 系统中，采用 TDD 模式，可以根据业务的需要灵活地调整上行和下行链路间的转换点，这样使双向业务成为了可能。对于非对称的无线多媒体通信，这是一种实现具有灵活资源管理的高速数据传输的方案之一。当用户的上行链路数据大于下行链路数据时，可以调整子帧中的转换点，使用户可以使用的时隙数增多，分配给该用户的 OFDM 符号数相应增加，满足用户高数据速率的需要；当用户的上行链路数据较少，请求低的数据速率时，调整子帧中的转换点，减少用户使用的时隙数，分配给该用户的 OFDM 符号数相应减少。正是由于这种分配给用户的 OFDM 符号数可变，使 OFDM-TDMA 多址方式可支持具有不同数据速率的多种业务。

OFDM-TDMA 系统的特点如下。

(1) OFDM-TDMA 系统能够避免窄带衰落，抗多径衰落能力较好。

(2) OFDM-TDMA 系统的信令开销很大程度上取决于是否采用滤除具有较低信噪比子载波的技术和自适应调制/编码技术，采用这些技术虽然可以改善性能，同时也增加了信令开销。

(3) OFDM-TDMA 系统在特定 OFDM 符号内将全部带宽分配给一个用户，这就不可避免地存在频率资源浪费、频带利用率较低和灵活性差等问题。

(4) OFDM-TDMA 系统要求相干检测，信道估计和均衡必不可少，因此，OFDM-TDMA



系统的多用户检测问题、信道估计和均衡问题将是未来研究的重点。

5.6.2　OFDMA

正交频分多址(OFDMA)技术是 OFDM-FDMA 的简写。OFDMA 为每个用户分配一个 OFDM 符号中的一个子载波或一组子载波，以子载波频率的不同来区分用户。OFDMA 接入方式与传统的 FDMA 方式类似之处是通过为每个用户提供部分不同的子载波来实现多用户接入，与 FDMA 不同之处在于 OFDMA 方式不需要在各个用户频率之间采用保护频段去区分不同的用户，这大大提高了系统的频率利用率。同时，基站通过调整子载波，可以根据用户的不同需求传输不同的速率。

给用户分配子载波有很多方法，使用最广泛的有两种：分组子载波和间隔扩展子载波，这两种方法的示意图如图 5.16 所示。分组子载波为每个用户分配一组相邻的子载波；而间隔扩展子载波为每个用户分配的子载波是有间隔的，也就是用户所使用的子载波扩展到整个系统带宽。分组子载波方法较简单，用户间干扰较小，但是受传输信道衰落的影响比较大；间隔扩展子载波方法通过频域扩展，增加频率分集，从而减少了信道衰落的影响，但其缺点是受用户干扰影响比较大，对同步要求较高。在 IEEE 802.16 的 OFDMA 模式中就采用了这种间隔扩展子载波分配方式。

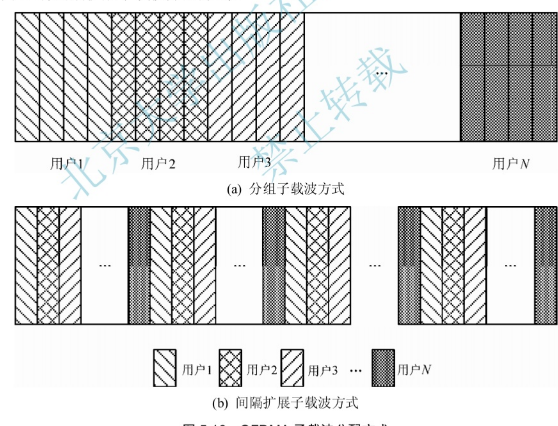
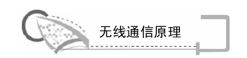


图 5.16　OFDMA 子载波分配方式

OFDMA 接入技术可以很容易地引入跳频技术，在每个时隙中，可以根据跳频图样来选择每个用户所使用的子载波频率，这种多址方式通常被称为 FH-OFDMA。FH-OFDMA 的优点是为小区内的多个用户设计正交跳频图样，可以很容易地消除小区内的干扰。由于



小区内的用户所经历的无线信道是不同的，对某个用户来说是最好的子载波，对其他用户很可能不是最好的。在 FH-OFDMA 的基础上，发送端如果知道每个用户的信道响应信息，就可以为每个用户分配信噪比高的子载波，这样，大部分用户可以分配到较好的子载波，从而获得多用户分集或位置分集，这种方法称为自适应子载波分配或自适应跳频。

OFDMA 多址方式的特点有以下几方面。

(1) OFDMA 可以灵活地适应带宽的要求，它通过简单地改变所使用的子载波数目就可以适应特定的传输带宽。

(2) OFDMA 系统可以不受小区内的干扰，这可以通过为小区内的多用户设计正交跳频图案来实现。

(3) OFDMA 可与动态信道分配技术相结合，以支持高速率的数据传输。

(4) OFDMA 可通过频域多用户调度和自适应编码技术，获得比 CDMA 更高的多用户分集增益。

5.6.3 多载波 CDMA

OFDM 和多址技术的结合能够允许多个用户同时共享有限的无线频谱，从而获得较高的系统容量。在这些多址技术中，CDMA 以其诸多的优点，并可提供比 FDMA 和 TDMA 更高的系统容量，成为第三代移动通信系统标准中采用的多址接入方式，将 OFDM 技术与传统的 CDMA 技术相结合的多址方式称为多载波 CDMA 技术。多载波 CDMA 不仅可以满足更多用户共享频率资源，而且同时可以减少码间干扰，提高系统性能。多载波 CDMA 方式主要有 MC-CDMA(Multicarrier CDMA)、MC-DS-CDMA(Multicarrier DS-CDMA)和 MT-CDMA(Multitone CDMA)3 种。其中 MC-CDMA 属于频域扩频，它将一个扩频序列的各个码片分配到不同的子载波上传输，从而可以获得明显的频率分集；MC-DS-CDMA 和 MT-CDMA 都属于时域扩频，它们的每一个子载波传输的都是一路 DS-CDMA 信号，当子载波数为 1 时，它们都退变为普通的 DS-CDMA。

时域扩频是各个码片分时占有同一段被展宽了的频谱，而频域扩频则是各个码片在不同的频段上占有同一段时间。若采用普通的 FDM 则频域扩频所占的总带宽约等于时域扩频，而采用 OFDM 则使频域扩频的频带利用率提高近一倍，从而仅用时域扩频一半的带宽就能实现相同的扩频增益，如图 5.17 所示。

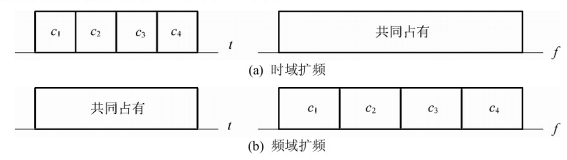


图 5.17 时域扩频与频域扩频的比较示意图

1. MC-CDMA

MC-CDMA 是最早提出的 OFDM 与 CDMA 结合的技术，MC-CDMA 采用频域扩频的方式，在频域使用给定的扩频序列将原始数据在不同的子载波上进行扩频。每一个信息符号由一个特定的扩频码进行扩频，然后将扩频后的每个码元符号调制到一个子载波上。图 5.18 为 MC-CDMA 基本原理图，图中 $a_i[k]$ 表示第 i 个用户输入数据序列中的一个调制符号，$\{c_{i,0},c_{i,1},\cdots,c_{i,N_c-1}\}$ 是与第 i 个用户对应的扩频码，$\{f_{i,0},f_{i,1},\cdots,f_{i,N_c-1}\}$ 是一组正交的子载波集，其频谱相互重叠。

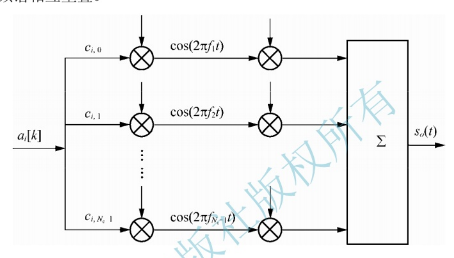


图 5.18　MC-CDMA 基本原理图

2. MC-DS-CDMA

MC-DS-CDMA 发送原理图和接收原理图分别如图 5.19 和图 5.20 所示。MC-DS-CDMA 是根据给定的扩频序列在时域扩频，MC-DS-CDMA 将发送数据序列先经过串/并转换变换成 N_c 路数据流并行输出，然后并行的每路数据由相同的扩频码序列进行扩频，使所得的各个子载波频谱在最小频率间隔的情况下保持正交，最后对这 N_c 路数据进行 OFDM 调制，调制后的信号进行求和发送。扩频后相邻子带之间有 1/2 的重叠，但保持正交关系，由于扩频后的信号带宽被限制在一个子带中，所以一般选择较短的扩频序列。

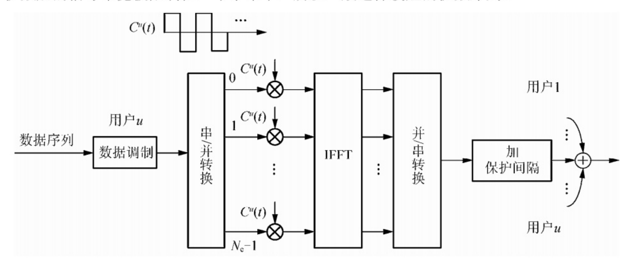


图 5.19　MC-DS-CDMA 发送原理框图

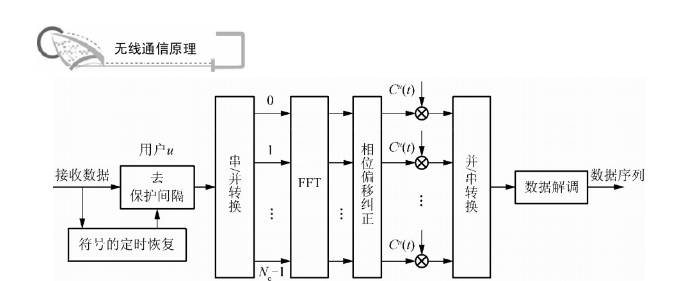


图 5.20 MC-DS-CDMA 接收原理框图

3. MT-CDMA

MT-CDMA 也是利用给定的扩频序列在时域扩频。MT-CDMA 信号调制示意图如图 5.21 所示。在 MT-CDMA 中，数据流先经过串/并转换后，再进行 OFDM 调制到不同的载波上，以形成 OFDM 信号，此时子载波之间有 1/2 的重叠，且满足正交性，然后再求和，最后对求和信号进行长扩频序列的时域扩频。扩频后，每个子载波的带宽得到扩展，即每个子载波的带宽变大，而相邻的子载波的间隔仍保持不变，因此在子载波之间有更多的重叠，但子载波之间已不再保持正交性关系。MT-CDMA 一般采用较长的扩频序列，比 DS-CDMA 能容纳更多的用户。

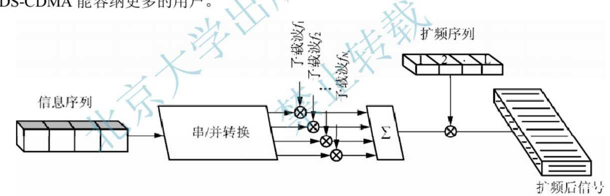


图 5.21 MT-CDMA 信号调制示意图

5.6.4 CS-OFDMA

码扩正交码分多址(CS-OFDMA)是 SCDMA 和 OFDMA 两种技术相融合得到的一种新的多址接入方式，这种多址技术结合了 CDMA 和 OFDMA 的优点，既能克服传统 CDMA 系统宽带数据传输时由扩展频谱引起的严重码间干扰，又能有效地对抗频率选择性衰落和相邻小区干扰，从而实现窄带语音和宽带数据的可靠传送。

CS-OFDMA 在进行 OFDM 调制之前，先对调制数据作码扩处理，多个数据信息通过伪随机正交码交织到多个相距一定间隔的频点上，从而使得每个频点的衰落不会严重影响数据信息的最后检测。CS-OFDMA 载波分配示意图如图 5.22 所示，带宽为 5MHz 的载波首先被划分成 5 个占用 1MHz 的子载波组，每个子载波组包含 128 个子载波。然后，每个



子载波组中的 128 个子载波又被均匀地分成 8 个子载波集合，每个集合包含 16 个连续的子载波。在调制过程中，每个用户调制信号的连续 N(1～8)个符号首先经过一个扩频系数为 8 的正交扩频调制，扩频调制后的 N 个信号相加产生 8 个码片信号，然后分别从这 8 个子载波集合中选取一个子载波(1/16)，把 8 个码片信号调制到对应的 8 个子载波上。这样就把每个符号的能量分散调制到了均匀分布在 1MHz 子载波组中的 8 个子载波上。

图 5.23 为 CS-OFDMA 系统的发射机结构，首先将宽带纵向分成许多窄带频点，然后再符号化，用正交 SCDMA 码道横向调制窄带频点，完成多调制信号的码扩处理，使每个符号在频域实现了扩频，然后进行 OFDM 调制后再发送。在频域进行码扩处理，可以对抗时域的瞬时干扰或时间选择性的深衰落，同时，码扩处理将每个符号的能量分到了整个信道的频带内，这在接收端可以实现频率分集接收。

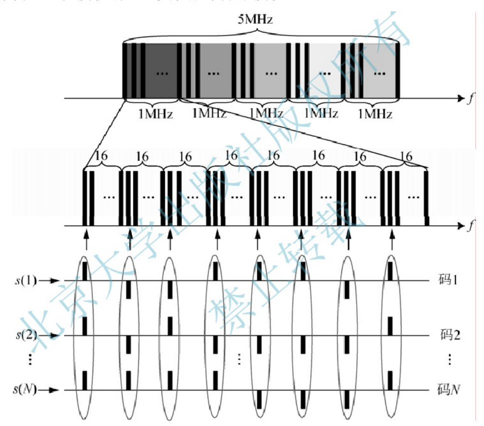


图 5.22　CS-OFDMA 载波分配示意图

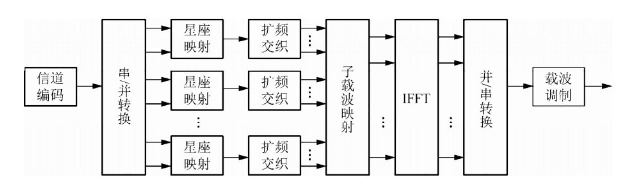


图 5.23　CS-OFDMA 系统的发射机结构

5.6.5 VSF-OFCDM

可变扩频因子-正交频分与码分复用多址技术是(VSF-OFCDM)是OFDM技术的又一种变型，主要用于下行传输。VSF-OFCDM 仍以 OFDM 作为基本的调制技术，它采用 OFDM 与 CDMA 结合的方法进行时频资源分配，其基本原理与 CS-OFDMA 较类似。VSF-OFCDM 系统发射机和接收机的结构分别如图 5.24 和图 5.25 所示。VSF-OFCDM 系统发射机主要处理过程：对输入数据信息进行编码后调制生成复制调制符号，然后进行串/并转换并送入交织器，将交织后的每一路输出复制成多路符号并乘以扩频序列后再进行码分复用，最后经过 IFFT 调制并添加循环前缀。VSF-OFCDM 接收机完成符号定时以及频域信道均衡后，之后的处理过程则为发射机的逆过程。

图 5.24 VSF-OFCDM 系统发射机结构图

图 5.25 VSF-OFCDM 系统接收机结构图

VSF-OFCDM 多址技术的特点如下。

(1) VSF-OFCDM 采用二维扩频结构，可达到很好的频率分集和时间分集的效果。

(2) 可以根据小区结构、信道负荷和传播环境，自适应地改变频域和时域扩频因子，

灵活地实现多小区和孤立小区的无缝接入。

(3) 在单小区情况下，令 SF=1，OFDMA 系统就是 OFDM 系统；在多小区情况下，令 SF>1，OFDMA 系统就等效于 MC-CDMA 系统。

(4) VSF-OFCDM 的发射机和接收机都包含了一个可变扩频因子控制器，当扩频因子改变时，相应的串/并转换、交织器和复制器的长度会相应地变化，这个控制器能够调整这些模块的长度变化，可以保证有足够的数据在同样多的子载波上发送。在接收端同样需要一个可变扩频因子控制器来调整对应的合并模块、解交织器和并/串转换长度。

5.7　随机多址技术

由于无线信道是一个公用信道，当有一个或多个分组信息需要发送时，用户接入信道，将分组信息发送出去。当多于一个用户试图同时发送分组时，分组在时间上将会重叠而发生冲突，因此必须通过多路接入通信协议来解决这样的冲突问题。随机多址接入就是使用自由竞争规则随机接入信道，采用网络协议形式实现的多址接入方式。常见的随机多址技术包括 ALOHA 协议、载波侦听多址(Carrier Sense Multiple Access，CSMA)协议和预约随机多址协议等。

5.7.1　ALOHA 随机多址方式

ALOHA(Additive Link On-line HAwaii)原本是夏威夷俚语，是用于对人到达或离开时致意的问候语。1968 年，夏威夷大学将解决夏威夷群岛之间数据通信的一项研究计划命名为 ALOHA。ALOHA 多址方式不同于前面介绍的时分、频分和码分的多址方式，它是以网络协议的形式来实现的，是一种最简单的数据分组传输协议。ALOHA 属于 ISO 七层协议中数据链路层的媒体访问控制 MAC 子层协议，用于共享信道的无线网络。从原理上看，共享信道的动态分配管理协议有两类：受控方式和随机方式。受控方式通常使用轮询和令牌两种方式，而随机方式自由竞争、冲突重发。若有两个以上用户同时发送，产生互相干扰而使发送不成功，这就产生了冲突，发生冲突以后，各个用户将随机等待一段时间以后再重新发送。

ALOHA 协议规定，任何一个用户随时有数据分组要发送，它就立刻接入信道进行发送。发送结束后，在相同的信道上或一个单独的反馈信道上等待应答。如果在一个给定的时间区间内，没有收到对方的认可应答，则重发刚发的数据分组。由于在同一信道上，多个用户独立随机地发送分组，就会出现多个分组发生冲突的情况，冲突的分组经过随机时延后重传。ALOHA 协议示意图如图 5.26(a)所示。从图中可以看出，要使当前分组传输成功，必须在当前分组到达时刻的前后各一个分组长度内没有其他用户的分组到达，即要保证到达的分组既没有整体冲突，也没有部分冲突，易损区间应为 2 倍的分组长度。

随机多址协议的主要性能指标有吞吐量(单位时间内平均成功传输的分组数)和每个分组的平均时延。为了改进 ALOHA 的性能，提出了改进的时隙 ALOHA 协议。改进的协议

将时间轴分成时隙，并且时隙大小要大于一个分组的长度，所有用户都同步在时隙开始时刻发送，称之为时隙 ALOHA 协议，如图 5.26(b)所示。与 ALOHA 协议相比，时隙 ALOHA 避免了部分冲突，并将易损区间减少为一个分组长度，提高了系统的吞吐量。

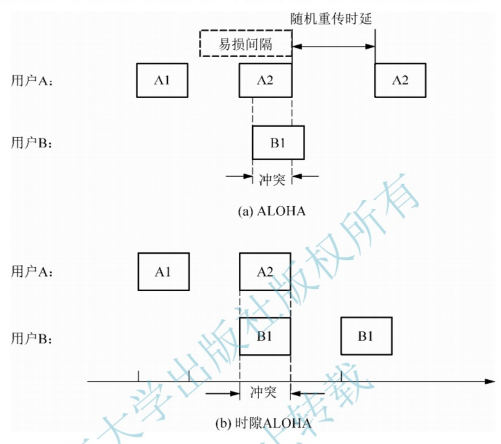


图 5.26　ALOHA 和时隙 ALOHA 协议示意图

5.7.2　CSMA 随机多址方式

ALOHA 协议在传输前并不监听信道，这意味着它对其他用户的情况并不了解，各个节点的发送是相互独立的。在载波侦听多址接入 CSMA 协议中，每个节点在发送前，首先要侦听信道是否有分组在传输。若没有检测到载波，信道处于空闲时，才可以发送；若信道忙，则按照设定的准则延迟发送。由于传送前侦听了信道，CSMA 随机多址方式减少了分组冲突的概率，进一步提高了系统的传输性能。CSMA 协议示意图如图 5.27 所示。在 CSMA 协议中，影响系统的两个主要参数是检测时延和传播时延。检测时延是指接收机判断信道是否空闲所需的时间。传播时延是指发送端发送数据开始到接收端收到数据所经历的时间。假定检测时延和传播时延之和为τ，如果某节点在 t 时刻开始发送一个分组，则在$t+\tau$时刻以后所有节点都会检测到信道忙，因此只要在$[t,t+\tau]$内没有其他用户发送，则该节点发送的分组将会成功传输。

存在 3 种基本的 CSMA：非坚持 CSMA、1-坚持 CSMA 和 p-坚持 CSMA。在非坚持 CSMA 中，若终端检测到信道忙，就不传送信号，并等待一个随机时延，按照一定的延迟



分布安排重传数据，时延后若检测到信道空闲，则传送数据，否则重复上述过程；在 1-坚持 CSMA 中，终端坚持继续检测信道直至信道空闲，一旦信道空闲则以概率 1 发送分组；在 p-坚持 CSMA 中，终端检测信道到信道空闲时，以概率 p 发送分组，以 $1-p$ 延迟发送。这 3 种 CSMA 的主要区别是检测信道空闲后分组传送方式不同。

CSMA 能够减少 ALOHA 协议的潜在冲突，终端在传输前检测信道状态决定分组传输或等待，但由于传播时延非零，而且终端之间没有协调程序，冲突仍有可能，但与 ALOHA 协议相比，CSMA 性能更优。

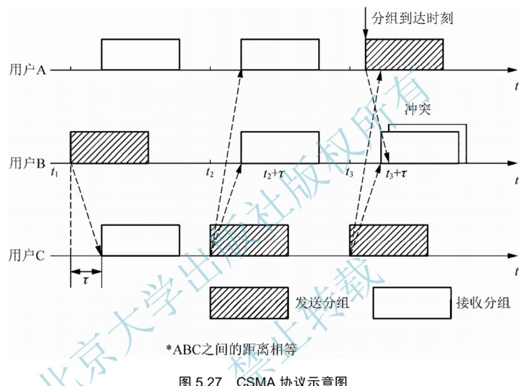


图 5.27　CSMA 协议示意图

5.7.3　预约随机多址方式

预约随机多址通常基于时分复用，即将时间轴分为重复的帧，每一帧分为若干时隙。当某用户有分组要发送时，可采用 ALOHA 方式在空闲时隙上进行预约。如果预约成功，它将无冲突地占用每一帧所预约的时隙，直至所有分组传输完毕。用于预约的时隙可以是一帧中固定的时隙，也可以是不固定的。预约时隙的大小可与信息传输时隙相同，也可以将一个时隙再分为若干个小时隙，每个小时隙供一个用户发送预约分组。

分组预约多址(Packet Radio Multiple Access，PRMA)是典型的预约随机多址方式，它是对 TDMA 的改进。PRMA 在 TDMA 帧结构的基础上为每一个话音突发在 TDMA 帧中预约一个时隙(TDMA 中一路话音固定占用一个时隙，而不管该话路是否有话音要传送)，预约的方法是当一个话音突发到达时，该节点在一帧中寻找空闲时隙，并在空闲时隙上发送该突发的第一个分组，如果传输成功，则它就预约了后续帧中对应的时隙，直至该突发传输结束。采用分组预约多址可以提高 TDMA 的性能。



5.7.4 其他随机多址方式

1. 忙音多址

忙音多址(BTMA)是解决秘密终端问题的一种技术，如两个终端可能处于第 3 个终端干扰的范围内，此时，系统频谱带宽分为两类信道，信息信道和忙音信道。当用户终端在信息信道上检测到信号时，它就接通忙音信道，当一用户终端检测到另一用户在信息信道上，则该用户终端就发送告警信号通知其他用户。忙音多址技术主要用于军事领域中，且移动台必须始终保持相互通信。

2. 数据侦听多址

数据侦听多址(DSMA)是一种全双工无线通信数据网络，前向信道(基站到移动台)将忙闲比特发送到每个数据帧，当忙闲比特变为空闲时，通知用户终端发送，用户一旦接收，基站就转换为忙比特状态。DSMA 用于高级移动电话业务、蜂窝数字分组数据和高级无线电数据信息业务。

3. 捕获分配分组接入

捕获分配分组接入(CDPA)以分组交换技术为基础并在所有小区中使用相同信道。ALOHA 重发机制用来解决不同小区传输之间的竞争问题，在一个小区内 CDPA 使用混合的预约轮询机理，由于传播时延小，这种机理容易综合不同类型的业务，CDPA 与编码方法相结合，并通过增加受传播条件和干扰变化控制的传输速率来提供恒定的吞吐量。CDPA 在每个小区使用相同的频道组，可以克服系统中的共信道冲突。

本 章 小 结

本章主要介绍了无线通信多址技术。多址技术的目的是多个用户共享信道、动态分配网络资源，在保证多用户之间通信质量的同时尽可能地降低系统的复杂度，并获得较高的系统容量。频道、时隙和码型是多址连接信道的 3 种主要形式，常见的多址接入方式有时分多址(TDMA)、频分多址(FDMA)、码分多址(CDMA)和空分多址(SDMA)。本章还介绍了正交频分多址(OFDMA)和随机多址技术等。FDMA 中是指各用户占用的频段；TDMA 中是指各用户占用的时隙；CDMA 中是指各用户占用的码型。SDMA 是通过空间的分割来区别不同用户，它利用无线的方向性波束将小区划分为不同的子空间来实现空间的正交隔离。正交频分复用(OFDM)是一种调制方式，它可以很容易地与传统的多址技术结合实现多用户接入服务，如 OFDM-TDMA、OFDMA 和多载波 CDMA 等。常见的随机多址协议包括 ALOHA 协议、载波侦听多址(CSMA)协议和预约随机多址协议。

习　　题

5.1 填空题

(1) 3 种基本的 CSMA 方式是__________、__________和__________。

(2) 多载波 CDMA 方式主要有__________、__________和__________。

5.2 简述多址方式的基本原理。简述 FDMA、TDMA、CDMA 和 SDMA 多址方式的基本原理。

5.3 FDMA、TDMA、CDMA 和 SDMA 多址方式分别有哪些特点？

5.4 什么是 m 序列？它有哪些特性？

5.5 什么是 Gold 序列？它有哪些特性？

5.6 正交频分复用(OFDM)调制方式可与传统的多址方式结合构成哪些新的多址接入技术？

5.7 随机多址方式有哪些？